GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 09:49:02 ; Search time 118.14 Seconds
                                    (without alignments)
                                    235.123 Million cell updates/sec

Title: US-09-596-746A-24
Perfect score: 1842
Sequence: 1 MACNNSGKDGNTSANSADES..........KNLSKAAKEMLTNSVKELTS 375
Scoring table: BLOSUM62
Searched: Gapop 10.0 , Gapext 0.5
                522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463
Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database : A_Geneseq_1101:*
    1: /SIDS8/gcgdata/geneseq/geneseqp/AA1980.DAT:*
    2: /SIDS8/gcgdata/geneseq/geneseqp/AA1981.DAT:*
    3: /SIDS8/gcgdata/geneseq/geneseqp/AA1982.DAT:*
    4: /SIDS8/gcgdata/geneseq/geneseqp/AA1983.DAT:*
    5: /SIDS8/gcgdata/geneseq/geneseqp/AA1984.DAT:*
    6: /SIDS8/gcgdata/geneseq/geneseqp/AA1985.DAT:*
    7: /SIDS8/gcgdata/geneseq/geneseqp/AA1986.DAT:*
    8: /SIDS8/gcgdata/geneseq/geneseqp/AA1987.DAT:*
    9: /SIDS8/gcgdata/geneseq/geneseqp/AA1988.DAT:*
   10: /SIDS8/gcgdata/geneseq/geneseqp/AA1989.DAT:*
   11: /SIDS8/gcgdata/geneseq/geneseqp/AA1990.DAT:*
   12: /SIDS8/gcgdata/geneseq/geneseqp/AA1991.DAT:*
   13: /SIDS8/gcgdata/geneseq/geneseqp/AA1992.DAT:*
   14: /SIDS8/gcgdata/geneseq/geneseqp/AA1993.DAT:*
   15: /SIDS8/gcgdata/geneseq/geneseqp/AA1994.DAT:*
   16: /SIDS8/gcgdata/geneseq/geneseqp/AA1995.DAT:*
   17: /SIDS8/gcgdata/geneseq/geneseqp/AA1996.DAT:*
   18: /SIDS8/gcgdata/geneseq/geneseqp/AA1997.DAT:*
   19: /SIDS8/gcgdata/geneseq/geneseqp/AA1998.DAT:*
   20: /SIDS8/gcgdata/geneseq/geneseqp/AA1999.DAT:*
   21: /SIDS8/gcgdata/geneseq/geneseqp/AA2000.DAT:*
   22: /SIDS8/gcgdata/geneseq/geneseqp/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1837	99.7	374	22	AAB62710	Borrelia sp chimer
2	1834	99.6	398	22	AAB62728	Borrelia sp chimer
3	1819.5	98.8	401	22	AAB62738	Borrelia sp chimer
4	1810.5	98.3	378	22	AAB62712	Borrelia sp chimer
5	1592.5	86.5	373	22	AAB62711	Borrelia sp chimer
6	1589.5	86.3	397	22	AAB62729	Borrelia sp chimer
7	1575	85.6	400	22	AAB62739	Borrelia sp chimer
8	1575	85.5	397	22	AAB62713	Borrelia sp chimer
9	1566	85.0	377	22	AAB62711	Borrelia sp chimer
10	1529	82.8	370	22	AAB62730	Borrelia sp chimer
11	1493	81.1	368	22	AAB62718	Borrelia sp chimer
12	1492	81.0	560	22	AAB62724	Borrelia sp chimer
13	1490	80.9	392	22	AAB62734	Borrelia sp chimer
14	1488	80.8	384	22	AAB62726	Borrelia sp chimer
15	1485	80.6	408	22	AAB62737	Borrelia sp chimer
16	1284.5	69.7	369	22	AAB62716	Borrelia sp chimer
17	1281.5	69.6	393	22	AAB62731	Borrelia sp chimer
18	1248.5	67.8	367	22	AAB62719	Borrelia sp chimer
19	1245.5	67.7	391	22	AAB62735	Borrelia sp chimer
20	1213.5	65.9	378	22	AAB62725	Borrelia sp chimer
21	1212.5	65.8	369	22	AAB62714	Borrelia sp chimer
22	1209.5	65.7	393	22	AAB62736	Borrelia sp chimer
23	1205.5	65.4	401	22	AAB62733	Borrelia sp chimer
24	1181.5	64.1	368	22	AAB62717	Borrelia sp chimer
25	1173.5	63.7	391	22	AAB62740	Borrelia sp chimer
26	1165	63.2	410	22	AAB62727	Borrelia sp chimer
27	1164	63.2	386	22	AAB62721	Borrelia B31 outer surface
28	964	52.3	466	16	AAR75740	Borrelia outer surface
29	928	50.4	192	22	AAB62721	Outer surface prot
30	928	50.4	210	16	AAR75727	Outer surface prot
31	928	50.4	210	16	AAR75727	B31 outer surface
32	927	50.3	587	16	AAR75746	B31 Osp-A/antigen
33	926	50.3	209	22	AAR60886	Borrelia IP2 OspC
34	920	49.9	209	22	AAB62720	Borrelia IP2 OspC
35	915	49.7	192	22	AAR75739	B31 outer surface
36	908	49.3	190	22	AAB62703	B31 outer surface
37	885	48.0	190	22	AAB62707	B burgdorferi ospC
38	879	47.7	192	16	AAR60889	B burgdorferi ospC
39	864	46.9	192	15	AAR62772	Borrelia B31 antig
40	852	46.3	176	15	AAR62775	Borrelia 297 antig
41	694.5	37.7	177	15	AAR62773	Borrelia 25015 ant
42	694.5	37.7	177	15	AAR60887	Borrelia 25015 Osp
43	678.5	36.8	193	22	AAB62709	B burgdorferi ospC
44	676	36.8	212	16	AAW11934	Outer surface prot
45	674.5	36.6	191	15	AAR60897	Borrelia VS461 ant

ALIGNMENTS

RESULT 1
AAB62710

ID   AAB62710 standard; protein; 374 AA.
XX
AC   AAB62710;
XX
DT   03-APR-2001 (first entry)
XX
DE   Borrelia sp chimeric ospC protein SEQ ID NO: 24.
XX
KW   Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
OS   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
XX
PN   WO200078966-A1.
XX
PD   28-DEC-2000.
XX
PF   19-JUN-2000; 2000WO-US016915.
XX
PR   18-JUN-1999; 99US-0140042.
XX
PA   (UYNY) UNIV NEW YORK STATE RES FOUND.
PA   (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI   Dattwyler RJ, Selnost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
DR   WPI; 2001-050113/06.
DR   N-PSDB; AAP29014.
XX
PT   Compositions of OspC polypeptides from strains of Borrelia which cause
PT   Lyme disease are used to immunize animals and detect immune responses
PT   to Lyme disease -

RESULT 2

```
XX
PS  Claim 43; Page 78-79; 160pp; English.
XX
CC  The present invention provides compositions comprising ospc proteins and
CC  chimeric ospc proteins from members of the Borrelia genus. These may be
CC  Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC  vaccines against Borrelia infection, which is spread by ticks and leads
CC  to Lyme disease.
XX
SQ  Sequence   374 AA;

Query Match             99.7%;  Score 1837;  DB 22;  Length 374;
Best Local Similarity  100.0%;  Pred. No. 3.2e-119;
Matches 374;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    2 ACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAI 61
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 acnnsgkdgntsansadesvkgpnlteiskkitdsnavllavkeveallssideiaakai 60

Qy   62 GKKIHQNNGLDTEYNHNGSLLAGAYAISTLIQOKLDGLKNEGLKEKIDAARKCSETFTNK 121
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 gkkihqnngldteynhngsllagayaistliqokldglkneglkekidaarkcsetftnk 121

Qy  122 LKEKIHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 181
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 lkekihtdlgkegvtdadakeailktngtktkgaeelgklfesvevlskaakemlansvke 180

Qy  182 LTSPVVAESPAMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVETLL 241
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 ltspvvaespamvnnsgkdgntsansadesvkgpnlteiskkitesnavvlavkevetll 240

Qy  242 TSIDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAEIEK 301
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 tsidelakaigkkikndvsldneadhngslisgaylisnlitkkisaikdsgelkaeiek 300

Qy  302 AKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLSKA 361
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 akkcseeftaklkgehtdlgkegvtddnakkailktnndktkgadeleklfesvknlska 360

Qy  362 AKEMLTNSVKELTS 375
         ||||||||||||||
Db  361 akemltnsvkelts 374
```

RESULT 2
AAB62728
AC AAB62728;
XX
DT 03-APR-2001 (first entry)
XX
DE AAB62728 standard; Protein; 398 AA.
XX
KW Borrelia sp chimeric ospc protein SEQ ID NO: 60.
XX
OS Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
OS Chimeric - Borrelia sp.
OS Chimeric - Borrelia sp.
XX
PN WO200078966-A1.
XX
PD 28-DEC-2000.
XX
PF 19-JUN-2000; 2000WO-US16915.
XX
PR 18-JUN-1999; 99US-0140042.
XX
PA (UYNY) UNIV NEW YORK STATE RES FOUND.
PA (BROO-) BROOK BIOTECHNOLOGIES INC.
PI Dattwyler RJ, Selnost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;

RESULT 3

```
DR  WPI: 2001-050113/06.
DR  N-PSDB; AAP29032.
XX
CC  The present invention provides compositions comprising ospc proteins and
CC  chimeric ospc proteins from members of the Borrelia genus. These may be
CC  Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC  vaccines against Borrelia infection, which is spread by ticks and leads
CC  to Lyme disease.
XX
PT  Compositions of OspC polypeptides from strains of Borrelia which cause
PT  Lyme disease are used to immunize animals and detect immune responses
PS  Claim 43; Page 123; 160pp; English.
XX
SQ  Sequence   398 AA;

Query Match             99.6%;  Score 1834;  DB 22;  Length 398;
Best Local Similarity   99.7%;  Pred. No. 5.6e-119;
Matches 373;  Conservative  1;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    2 ACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAI 61
         :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   25 scnnsgkdgntsansadesvkgpnlteiskkitdsnavllavkeveallssideiaakai 84

Qy   62 GKKIHQNNGLDTEYNHNGSLLAGAYAISTLIQOKLDGLKNEGLKEKIDAARKCSETFTNK 121
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   85 gkkihqnngldteynhngsllagayaistliqokldglkneglkekidaarkcsetftnk 144

Qy  122 LKEKIHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 181
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  145 lkekihtdlgkegvtdadakeailktngtktkgaeelgklfesvevlskaakemlansvke 204

Qy  182 LTSPVVAESPAMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVETLL 241
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  205 ltspvvaespamvnnsgkdgntsansadesvkgpnlteiskkitesnavvlavkevetll 264

Qy  242 TSIDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAEIEK 301
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  265 tsidelakaigkkikndvsldneadhngslisgaylisnlitkkisaikdsgelkaeiek 324

Qy  302 AKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLSKA 361
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  325 akkcseeftaklkgehtdlgkegvtddnakkailktnndktkgadeleklfesvknlska 384

Qy  362 AKEMLTNSVKELTS 375
         ||||||||||||||
Db  385 akemltnsvkelts 398
```

RESULT 3
AAB62738
ID AAB62738 standard; Protein; 401 AA.
XX
AC AAB62738;
XX
DT 03-APR-2001 (first entry)
XX
DE AAB62738 standard; Protein; 401 AA.
XX
KW Borrelia sp chimeric ospc protein SEQ ID NO: 80.
XX
OS Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
OS Chimeric - Borrelia sp.
OS Chimeric - Borrelia sp.
XX
PN WO200078966-A1.
XX
PD 28-DEC-2000.
XX
PF 19-JUN-2000; 2000WO-US16915.
PR 18-JUN-1999; 99US-0140042.

Left column

```
XX
PA   (UYNY ) UNIV NEW YORK STATE RES FOUND.
PA   (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI   Dattwyler RJ, Seinost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
DR   WPI; 2001-050113/06.
DR   N-PSDB; AAF29042.
XX
PA
PA
XX
PT   Compositions of OspC polypeptides from strains of Borrelia which cause
PT   Lyme disease are used to immunize animals and detect immune responses
PT   to Lyme disease -
XX
CC   The present invention provides compositions comprising ospc proteins and
CC   chimeric ospc proteins from members of the Borrelia genus. These may be
CC   Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC   vaccines against Borrelia infection, which is spread by ticks and leads
CC   to Lyme disease.
XX
PS   Claim 43; Page 150-151; 160pp; English.
XX

Query Match             98.8%;  Score 1819.5;  DB 22;  Length 401;
Best Local Similarity   98.7%;  Pred. No. 5.6e-118;
Matches  372;  Conservative  2;  Mismatches  0;  Indels  3;  Gaps  1;

Sequence   401 AA;

Qy      2 ACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAI 61
Db     25 scnnsgkdgntsansadesvkgpnlteiskkitdsnavllavkeveallssideiaakai 84
Qy     62 GKKIHONNGLDTEYNHNGSLLAGAYATSTLIKOKLDGLKNEGLKEKIDAAKKCSETFPNK 121
Db     85 gkkihqnngldteynhngsllagayaistllikqkldglknegikeidaakkcsetftnk 144
Qy    122 LKEKHTDLGKECGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 181
Db    145 lkekhtdlgkegvtddadakeailktngtktkgaeelgklfesvevlskaakemlansvke 204
Qy    182 LTSPVVAES---PAMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVE 238
Db    205 ltspvvaespkpsmvnnsgkdgntsansadesvkgpnlteiskkitesnavvlavkeve 264
Qy    239 TLLTSIDELAKAIGKKIKKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAE 298
Db    265 tlltsidelakaigkkikndvsldneadhngslisgaylisnlitkkisaikdsgelkae 324
Qy    299 IEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNL 358
Db    325 iekakkcseeftaklkgehtdlgkegvtddnakkailktnndktkgadeleklfesvknl 384
Qy    359 SKAAKEMLTNSVKELTS 375
Db    385 skaakemltnsvkelts 401
```

```
RESULT  4
        AAB62712
ID   AAB62712 standard; Protein; 378 AA.
XX
AC   AAB62712;
XX
DT   03-APR-2001  (first entry)
XX
DE   Borrelia sp chimeric ospc protein SEQ ID NO: 28.
XX
KW   Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
XX
PN   WO200078966-A1.
XX
```

Right column

```
XX
PD   28-DEC-2000.
XX
PF   19-JUN-2000; 2000WO-US16915.
XX
PR   18-JUN-1999; 99US-0140042.
XX
PA   (UYNY ) UNIV NEW YORK STATE RES FOUND.
PA   (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI   Dattwyler RJ, Seinost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
DR   WPI; 2001-050113/06.
DR   N-PSDB; AAF29016.
XX
PT   Compositions of OspC polypeptides from strains of Borrelia which cause
PT   Lyme disease are used to immunize animals and detect immune responses
PT   to Lyme disease -
XX
CC   The present invention provides compositions comprising ospc proteins and
CC   chimeric ospc proteins from members of the Borrelia genus. These may be
CC   Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC   vaccines against Borrelia infection, which is spread by ticks and leads
CC   to Lyme disease.
XX
PS   Claim 43; Page 83-84; 160pp; English.
XX

Query Match             98.3%;  Score 1810.5;  DB 22;  Length 378;
Best Local Similarity   98.1%;  Pred. No. 2.2e-117;
Matches  371;  Conservative  2;  Mismatches  2;  Indels  3;  Gaps  1;

SQ   Sequence   378 AA;

Qy      1 MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA 60
Db      1 macnnsgkdgntsansadesvkgpnlteiskkitdsnavllavkeveallssideiaaka 60
Qy     61 IGKKIHONNGLDTEYNHNGSLLAGAYAISTLIKOKLDGLKNEGLKEKIDAAKKCSETFTN 120
Db     61 igkkihqnngldtennhngsllagayaistllikqkldglknegikeidaakkcsetftn 120
Qy    121 KLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 180
Db    121 klkekhtdlgkegvtddadakeailkangtktkgaeelgklfesvevlskaakemlansvk 180
Qy    181 ELTSPVVAES---PAMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEV 237
Db    181 eltspvvaespkpsmvnnsgkdgntsansadesvkgpnlteiskkitesnavvlavkev 240
Qy    238 ETLLTSIDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKA 297
Db    241 etlltsidelakaigkkikndvsldneadhngslisgaylisnlitkkisaikdsgelka 300
Qy    298 EIEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKN 357
Db    301 eiekakkcseeftaklkgehtdlgkegvtddnakkailktnndktkgadeleklfesvkn 360
Qy    358 LSKAAKEMLTNSVKELTS 375
Db    361 lskaakemltnsvkelts 378
```

```
RESULT  5
        AAB62711
ID   AAB62711 standard; Protein; 373 AA.
XX
AC   AAB62711;
XX
DT   03-APR-2001  (first entry)
XX
DE   Borrelia sp chimeric ospc protein SEQ ID NO: 26.
XX
```

KW Borrelia; OspC; Lyme disease; vaccine; chimeric protein; tick.
OS Chimeric - Borrelia sp.
OS Chimeric - Borrelia sp.
XX
XX
PN WO200078966-A1.
XX
PD 28-DEC-2000.
XX
PF 19-JUN-2000; 2000WO-US16915.
XX
PR 18-JUN-1999; 99US-0140042.
XX
PA (UYNY) UNIV NEW YORK STATE RES FOUND.
PA (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI Dattwyler RJ, Seinost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
PR WPI: 2001-050113/06;
   N-PSDB; AAF29015.
XX
XX
CC The present invention provides compositions comprising ospc proteins and
CC chimeric ospc proteins from members of the Borrelia genus. These may be
CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC vaccines against Borrelia infection, which is spread by ticks and leads
CC to Lyme disease.
XX
PS Claim 43; Page 81; 160pp; English.
XX

SQ Sequence 373 AA;

Query Match 87.7%; Score 1592.5; DB 22; Length 373;
Best Local Similarity 86.5%; Pred. No. 2.4e-102;
Matches 327; Conservative 20; Mismatches 25; Indels 1; Gaps 1;

```
Qy   2 ACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAI  61
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 acnnsgkdgntsansadesvkgpnlteiskkitdsnavllavkeveallssideiaakai  60

Qy  62 GKKIHONNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNK 121
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  61 gkkihqnngldteynhngsllagayaistlikqkldglkneglkekidaakkcsetftnk 120

Qy 122 LKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 181
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 lkekhtdlgkegvtdadakeailktngtktkgaeelgklfesvevlskaakemlansvke 180

Qy 182 LTSPVVAESPAMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVETLL 241
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||
Db 181 ltspvvaespamvnnsgkdgntsansadesvkgpnlteiskkitesnavvlavkeietll 240

Qy 242 TSIDELA-KAIGKKIKNDVSLDNEADHNGSLISGAYLIISNLITKKISAIKDSGELKAEIE 300
        ::|||| :||||||:  :  :|: |||:||| |||||:  :|:| ||:|::||| :||
Db 241 asidelatkaigkkiqgnglqlaveaghngtllagayaistlikqkldglkneklkekie 300

Qy 301 KAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLSK 360
        :|||||::|| || |||| |:| |::|::|||| | |::| || |:| |||| |:||| |
Db 301 nakkcsedftkklegehaqlgienvtdenakkailtdaakdkgaaeleklfavenlak   360

Qy 361 AAKEMLTNSVKEL 373
        ||||||:||||||
Db 361 aakemlansvkel 373
```

RESULT 6
ID AAB62729 standard; Protein; 397 AA.
XX
AC AAB62729;
XX
DT 03-APR-2001 (first entry)
XX
DE Borrelia sp chimeric ospc protein SEQ ID NO: 62.
XX
KW Borrelia; OspC; Lyme disease; vaccine; chimeric protein; tick.
OS Chimeric - Borrelia sp.
OS Chimeric - Borrelia sp.
XX
PN WO200078966-A1.
XX
PD 28-DEC-2000.
XX
PF 19-JUN-2000; 2000WO-US16915.
XX
PR 18-JUN-1999; 99US-0140042.
XX
PA (UYNY) UNIV NEW YORK STATE RES FOUND.
PA (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI Dattwyler RJ, Seinost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
DR WPI: 2001-050113/06;
DR N-PSDB; AAF29033.
XX
CC The present invention provides compositions comprising ospc proteins and
CC chimeric ospc proteins from members of the Borrelia genus. These may be
CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC vaccines against Borrelia infection, which is spread by ticks and leads
CC to Lyme disease.
XX
PS Claim 43; Page 125-126; 160pp; English.
XX

SQ Sequence 397 AA;

Query Match 86.3%; Score 1589.5; DB 22; Length 397;
Best Local Similarity 87.4%; Pred. No. 4.2e-102;
Matches 326; Conservative 21; Mismatches 25; Indels 1; Gaps 1;

```
Qy   2 ACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAI  61
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  25 scnnsgkdgntsansadesvkgpnlteiskkitdsnavllavkeveallssideiaakai  84

Qy  62 GKKIHONNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNK 121
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  85 gkkihqnngldteynhngsllagayaistlikqkldglkneglkekidaakkcsetftnk 144

Qy 122 LKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 181
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 145 lkekhtdlgkegvtdadakeailktngtktkgaeelgklfesvevlskaakemlansvke 204

Qy 182 LTSPVVAESPAMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVETLL 241
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||
Db 205 ltspvvaespamvnnsgkdgntsansadesvkgpnlteiskkitesnavvlavkeietll 264

Qy 242 TSIDELA-KAIGKKIKNDVSLDNEADHNGSLISGAYLIISNLITKKISAIKDSGELKAEIE 300
        ::|||| :||||||:  :  :|: |||:||| |||||:  :|:| ||:|::||| :||
Db 265 asidelatkaigkkiqgnglqlaveaghngtllagayaistliitqkldglknseklkekie 324

Qy 301 KAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLSK 360
        :|||||::|| || |||| |:| |::|::|||| | |::| || |:| |||| |:||| :|
Db 325 nakkcsedftkklegehaqlgienvtdenakkailtdaakdkgaaeleklfavenlak   384

Qy 361 AAKEMLTNSVKEL 373
        ||||||:||||||
Db 385 aakemlansvkel 397
```

RESULT 7

AAB62739

ID   AAB62739 standard; Protein; 400 AA.
XX
AC   AAB62739;
XX
DT   03-APR-2001 (first entry)
XX
DE   Borrelia sp chimeric ospC protein SEQ ID NO: 82.
XX
KW   Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
XX
PN   WO200078966-A1.
XX
PD   28-DEC-2000.
XX
PR   19-JUN-2000; 2000WO-US16915.
XX
PR   18-JUN-1999;   99US-0140042.
XX
PA   (UYNY) UNIV NEW YORK STATE RES FOUND.
PA   (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI   Dattwyler RJ, Seinost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
PT   Claim 43; Page 153; 160pp; English.
XX
PF   18-JUN-1999;   99US-0140042.
XX
DR   WPI; 2001-050113/06.
DR   N-PSDB; AAF29043.
XX
CC   The present invention provides compositions comprising ospC proteins and
CC   chimeric ospC proteins from members of the Borrelia genus. These may be
CC   Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC   vaccines against Borrelia infection, which is spread by ticks and leads
CC   to Lyme disease.
XX
PS   Sequence   400 AA;
XX
SQ

Query Match   85.5%; Score 1575; DB 22; Length 400;
Best Local Similarity 86.4%; Pred. No. 4.3e-101;
Matches 325; Conservative 22; Mismatches 25; Indels   4; Gaps   2;

Qy   2 ACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAI 61
            :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   25 scnnsgkdgntsansadesvkgpnlteiskkitdsnavllavkeveallssideiaaka1 84

Qy   62 GKKIHONNGLDTEYNHNGSLLAGAYAISTLLIKOKLDGLKNEGLKEKIDAAKKCSETFTNK 121
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   85 gkkihqnngldteynhngsllagayaistllikqkldglkneglkekidaakkcsetftnk 144

Qy   122 LKEKEHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 181
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   145 lkekehtdlgkegvtdadakeailktngtktkgaeelgklfesvevlskaakemlansvke 204

Qy   182 LTSPVVAES---PAMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVLLAVKEVE 238
           |:|||||||| :|||||||||||||||||||||||||||||||||||||||||||:|
Db   205 ltspvvaespkkpsmvnnsgkdgntsansadesvkgpnlteiskkitesnavllavkeie 264

Qy   239 TLLTSIDELA-KAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITTKKISAIKDSGELKA 297
           ||| |||||| |||||||||||||||||||||||||||||||||||||||||||||:||
Db   265 tllasidelatkalgkkiqgnglaveaghnglliagayt1isklltqkldglknseklke 324

Qy   298 EIEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKN 357
           |||||:||||||||||||||||||||||||||||||||||||||||||||||||||:||
Db   325 kienakkcsedftkklegehaqlgienvtdenakkailtdaakdkgaaeleklfkaven 384

───

RESULT 8

AAB62713

ID   AAB62713 standard; Protein; 377 AA.
XX
AC   AAB62713;
XX
DT   03-APR-2001 (first entry)
XX
DE   Borrelia sp chimeric ospC protein SEQ ID NO: 30.
XX
KW   Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
XX
PN   WO200078966-A1.
XX
PD   28-DEC-2000.
XX
PR   19-JUN-2000; 2000WO-US16915.
XX
PR   18-JUN-1999;   99US-0140042.
XX
PA   (UYNY) UNIV NEW YORK STATE RES FOUND.
PA   (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI   Dattwyler RJ, Seinost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
PT   Claim 43; Page 86-87; 160pp; English.
XX
DR   WPI; 2001-050113/06.
DR   N-PSDB; AAF29017.
XX
CC   The present invention provides compositions comprising ospC proteins and
CC   chimeric ospC proteins from members of the Borrelia genus. These may be
CC   Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC   vaccines against Borrelia infection, which is spread by ticks and leads
CC   to Lyme disease.
XX
SQ   Sequence   377 AA;

Query Match   85.0%; Score 1566; DB 22; Length 377;
Best Local Similarity 85.9%; Pred. No. 1.7e-100;
Matches 324; Conservative 22; Mismatches 27; Indels   4; Gaps   2;

Qy   1 MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA 60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 macnnsgkdgntsansadesvkgpnlteiskkitdsnavllavkeveallssideiaaka 60

Qy   61 IGKKIHONNGLDTEYNHNGSLLAGAYAISTLLIKQKLDGLKNEGLKEKIDAAKKCSETFTN 120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 igkkihqnngldtennhngsllagayaistllikqkldglkneglkekidaakkcsetftn 120

Qy   121 KLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 klkekhtdlgkegvtdadakeailkangtktkgaeelgklfesvevlskaakemlansvk 180

Qy   181 ELTSPVVAES---PAMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVLLAVKEV 237
           |:|||||||| :|||||||||||||||||||||||||||||||||||||||||:|
Db   181 eltspvvaespkkpsmvnnsgkdgntsansadesvkgpnlteiskkitesnavllavkei 240

```
Qy   238  ETLLTSIDELA-KAIGKKIKNDVSLDNEADHNGSLLISGAYLISNLITKKISAIKDSGELK  296
          ||||:|:|||| :||:|| ||::|:|| ::|:|:||:|| |||| :|:||:|:|||:||
Db   241  etllasidelatkaigkiqcngglaveeaghngtllagaytiskliqtkidglknsekik  300

Qy   297  AEIEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVK  356
          :::.:||||||:||.||.|||:.|||:.|||:||||||:.|:.|.||||:||||||:.||
Db   301  ekienakkcsedftkklegehaglgienvtdenakkallitdaakdkgaaeleklfkave  360

Qy   357  NLSKAAKEMLTNSVKEL  373
          ||:|||||||:||||||
Db   361  nlakaakemlansvkel  377
```

```
RESULT   9

ID   AAB62715 standard; Protein; 370 AA.
XX
AC   AAB62715;
XX
DT   03-APR-2001  (first entry)
XX
DE   Borrelia sp chimeric ospC protein SEQ ID NO: 34.
XX
KW   Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
XX
PN   WO200078966-A1.
XX
PD   28-DEC-2000.
XX
PF   19-JUN-2000; 2000WO-US16915.
XX
PR   18-JUN-1999;  99US-0140042.
XX
PA   (UINY ) UNIV NEW YORK STATE RES FOUND.
PA   (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI   Dattwyler RJ, Selnost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
DR   WPI; 2001-050113/06.
DR   N-PSDB; AAF29019.
XX
PT   Compositions of OspC polypeptides from strains of Borrelia which cause
PT   Lyme disease are used to immunize animals and detect immune responses
PT   to Lyme disease.
XX
PS   Claim 43; Page 91-92; 160pp; English.
XX
CC   The present invention provides compositions comprising ospC proteins and
CC   chimeric ospC proteins from members of the Borrelia genus. These may be
CC   Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC   vaccines against Borrelia infection, which is spread by ticks and leads
CC   to Lyme disease.
XX
SQ   Sequence  370 AA;

Query Match             83.0%;  Score 1529;  DB 22;  Length 370;
Best Local Similarity   87.0%;  Pred. No. 5.8e-98;
Matches  327;  Conservative   9;  Mismatches  32;  Indels   8;  Gaps   4;

Qy    2  ACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAI   61
         |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db    1  acnnsgkdgntsansadesvkgpnlteiskkitdsnavllavkeveallssidel-akai   59

Qy   62  GKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLK-NEGLKEKIDAAKKCSETFIN  120
         |||  ||| ||| :||||||||| | |||||| : ||||  ||||||||:||||||:|:
Db   60  gkkikndgldneanrnesllagaytistltltqklsklngseglkekiaaakkcseefst  119

Qy  121  KLKEKHTDLGKEGVTDADAKEAILKTNGT-KTKGAEELGKLFESVEVLSKAAKEMLANSV  179
         |||::| :||:|:||| :|:|||||:|    | ||.|||.||:.|:: |||||||||||||
Db  121  klkdnhaqlgigqvtdenakkallkanaagkdkgveeleklsgslesiskaakemlansv  179
```

```
RESULT   10

ID   AAB62730 standard; Protein; 394 AA.
XX
AC   AAB62730;
XX
DT   03-APR-2001  (first entry)
XX
DE   Borrelia sp chimeric ospC protein SEQ ID NO: 64.
XX
KW   Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
XX
PN   WO200078966-A1.
XX
PD   28-DEC-2000.
XX
PR   18-JUN-1999;  99US-0140042.
XX
PF   19-JUN-2000; 2000WO-US16915.
XX
PA   (UINY ) UNIV NEW YORK STATE RES FOUND.
PA   (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI   Dattwyler RJ, Selnost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
DR   WPI; 2001-050113/06.
DR   N-PSDB; AAF29034.
XX
PT   Compositions of OspC polypeptides from strains of Borrelia which cause
PT   Lyme disease are used to immunize animals and detect immune responses
PT   to Lyme disease.
XX
PS   Claim 43; Page 128-129; 160pp; English.
XX
CC   The present invention provides compositions comprising ospC proteins and
CC   chimeric ospC proteins from members of the Borrelia genus. These may be
CC   Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC   vaccines against Borrelia infection, which is spread by ticks and leads
CC   to Lyme disease.
XX
SQ   Sequence  394 AA;

Query Match             82.8%;  Score 1526;  DB 22;  Length 394;
Best Local Similarity   86.7%;  Pred. No. 1e-97;
Matches  326;  Conservative  10;  Mismatches  32;  Indels   8;  Gaps   4;

Qy    2  ACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAI   61
         |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db   25  scnnsgkdgntsansadesvkgpnlteiskkitdsnavllavkeveallssidel-akai   83

Db  120  klkdnhaqlgigqvtdenakkallkanaagkdkgveeleklsgslesiskaakemlansv  179

Qy  180  KELTSPVVAEASPAMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVET  239
         |||:||||       :||||||||||||||||||||||||||||||||||||||||||||
Db  180  keltspvvhg-----nnsgkdgntsansadesvkgpnlteiskkitesnavvlavkevet  234

Qy  240  LLTSIDELAKAIGKKIKNDVSLDNEADHNGSLLISGAYLISNLITKKISAIKDSGELKAEI  299
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  235  lltsidelakaigkkikndvsldneadhngsllisgaylisnlitkkisaikdsgelkaei  294

Qy  300  EKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLS  359
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  295  ekakkcseeftaklkgehtdlgkegvtddnakkailktnndktkgadeleklfesvknls  354

Qy  360  KAAKEMLTNSVKELTS  375
         ||||||||||||||||
Db  355  kaakemltnsvkelts  370
```

```
RESULT 11
ID   AAB62718 standard; Protein; 368 AA.
XX
AC   AAB62718;
XX
DT   03-APR-2001  (first entry)
XX
DE   Borrelia sp chimeric ospC protein SEQ ID NO: 40.
XX
KW   Borrelia; ospC; Lyme disease; vaccine; chimeric; protein; tick.
XX
OS   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
XX
PN   WO200078966-A1.
XX
PD   28-DEC-2000.
XX
PF   19-JUN-2000; 2000WO-US16915.
XX
PR   18-JUN-1999;  99US-0140042.
XX
PA   (UNY ) UNIV NEW YORK STATE RES FOUND.
PA   (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI   Dattwyler RJ, Selnost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
DR   WPI; 2001-050113/06.
DR   N-PSDB; AAF29022.
XX
PT   Claim 43; Page 99-100; 160pp; English.
XX
PT   Compositions of OspC polypeptides from strains of Borrelia which cause
PT   Lyme disease are used to immunize animals and detect immune responses
PT   to Lyme disease -
XX
CC   The present invention provides compositions comprising ospC proteins and
CC   chimeric ospC proteins from members of the Borrelia genus. These may be
CC   Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC   vaccines against Borrelia infection, which is spread by ticks and leads
CC   to Lyme disease.
XX
SQ   Sequence  368 AA;

  Query Match             81.18;  Score 1493;  DB 22;  Length 368;

Qy        62 GKKIHQNNGLDTEYNINGSLLAGAYAISTLIKQKLDGLK-NEGLKEKIDAAKKCSEFFTN 120
              ||| |||| ||| || |||||||||| ||| ||| ||||  |||||||| ||||||| :
Db        84 gkkikndgsldneanrnesllagaytistllcqklskingseglkekiaaakkcseefst 143

Qy       121 KLKEKHFDLGKEGVTDADAKEAILKTNGT-KTKGAEELGKLFESVEVLSKAAKEMLANSV 179
              ||| :| ||| ||||| ||||| || ::|  :|||| ||| |||| |||||||||||||
Db       144 klkdnhaqlglqgvtdenakkailkanaagkdkgveeleklsgsleslskaakemlansv 203

Qy       180 KELTSPVVAESPAMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVET 239
              ||||||||        |||||||||||||||||||||||||||||||||||||||||||
Db       204 keltspvvhg-----nnsgkdgntsansadesvkgpnlteiskkitesnavvlavkevet 258

Qy       240 LLTSIDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAEI 299
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       259 lltsidelakaigkkikndvsldneadhngslisgaylisnlitkkisaikdsgelkaei 318

Qy       300 EKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLS 359
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       319 ekakkcseeftaklkgehtdlgkegvtddnakkailktnndktkgadeleklfesvknls 378

Qy       360 KAAKEMLTNSVKELTS 375
              ||||||||||||||||
Db       379 kaakemltnsvkelts 394
```

```
RESULT 12
ID   AAB62724 standard; Protein; 560 AA.
XX
AC   AAB62724;
XX
DT   03-APR-2001  (first entry)
XX
DE   Borrelia sp chimeric ospC protein SEQ ID NO: 52.
XX
KW   Borrelia; ospC; Lyme disease; vaccine; chimeric; protein; tick.
XX
OS   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
XX
PN   WO200078966-A1.
XX
PD   28-DEC-2000.
XX
PF   19-JUN-2000; 2000WO-US16915.
XX
PR   18-JUN-1999;  99US-0140042.
XX
PA   (UNY ) UNIV NEW YORK STATE RES FOUND.
PA   (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI   Dattwyler RJ, Selnost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
DR   WPI; 2001-050113/06.
DR   N-PSDB; AAF29028.
XX
PT   Claim 43; Page 112-113; 160pp; English.
XX
PT   Compositions of OspC polypeptides from strains of Borrelia which cause
PT   Lyme disease are used to immunize animals and detect immune responses
PT   to Lyme disease -
XX
CC   The present invention provides compositions comprising ospC proteins and
CC   chimeric ospC proteins from members of the Borrelia genus. These may be
CC   Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC   vaccines against Borrelia infection, which is spread by ticks and leads

  Best Local Similarity   83.4%;
  Matches  312;  Conservative   22;  Mismatches   34;  Indels    6;  Gaps    2;
  Pred. No. 1.8e-95;

Qy         2 ACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAI 61
              |||||||||:|||||||||||||||||||||||:||||:|||||||:|||:||:||:||
Db         1 acnnsgkdgnasansadesvkgpnlteiskkitesnavllavkevetllasidelatkai 60

Qy        62 GKKIHQNNGLDTEYNINGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSEFFTNK 121
              |||: ||||| |   ||| ||:||||| :| || :  |||||:|||||: ||:||:|||
Db        61 gkki-gnngleangsknitsllsgayaisdtlaeklnvlkneelkekidtakqcsteftnk 119

Qy       122 LKEKHFDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 181
              | | | :||   :|  ||:|| ||:: :  |||||||| |: ||:||||| ::|:||||
Db       120 lksehavlgldnltddnaqtrailkhankdgaaeleklfkavenlskaaqqtlknavke 179

Qy       182 LTSPVVAESPAMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVETLL 241
              ||||:|        |||||||||||||||||||||||||||||||||||||||||||||
Db       180 ltspivhg-----nnsgkdgntsansadesvkgpnlteiskkitesnavvlavkevetll 234

Qy       242 TSIDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAEIEK 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       235 tsidelakaigkkikndvsldneadhngslisgaylisnlitkkisaikdsgelkaeiek 294

Qy       302 AKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLSKA 361
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       295 akkcseeftaklkgehtdlgkegvtddnakkailktnndktkgadeleklfesvknlska 354

Qy       362 AKEMLTNSVKELTS 375
              ||||||||||||||
Db       355 akemltnsvkelts 368
```

RESULT 13

```
CC   to Lyme disease.
XX
SQ   Sequence 560 AA;

Query Match             Best Local Similarity 81.0%;  Score 1492; DB 22; Length 560;
                        Matches 313; Conservative 20; Mismatches 40; Indels 2; Gaps 2;

Qy     1 MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 macnnsgkdgntsansadesvkgpnlteiskkitdsnavllavkeveallssideiaaka 60

Qy    61 IGKIHQNNGLDTEYNHNGSLLAGAYAISTLLIKQKLDGLKNEGLKEKIDAAKKCSETFTN 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 igkihqnngldteynhngsllagayaistllikqkldglkneglkekidaakkcsetftn 120

Qy   121 KLREKHTDLGKEGVTDADAKEAILKTNGTKTKTGAEELGKLFESVEVLSKRAAKEMLANSVK 180
          |||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||
Db   121 klkekhtdlgkegvtdadakeailktngtktktgaeelgklfesvevlskaakemlansvk 180

Qy   181 ELTSPVVAESPAMVNNSGKDGNT-SANSADESVKGPNLTEISKKITESNAVVLAVKEVET 239
          :| ||||| :||| :||||||| | |||||||:||||||||||||||||||||||||||||
Db   181 eltspvvaespamvnnsgkggdsastnpadesakgpnlteiskkitdsnafvlavkevet 240

Qy   240 LLTSIDELA-KAIGKIKINDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKRAE 298
          :| |||||| ||||:|:|:|||||||||||||||||||||||||||||||||||||||||
Db   241 lvisidelaktaigqkidnnglaalnngngsllagayaistlitekiskiknleelkte 300

Qy   299 IEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKRAILKTNNDKTKGADELEKLFESVKNL 358
          ||||||||||||| || |:|||||:|| |::|| ||||| |||:||||||||||||||||
Db   301 lakakkcseeftnkiksghadlgkqdatddhakaailkthattdkgakefkdlfesvegl 360

Qy   359 SKAAKEMLTNSVKEL 373
          ||||: ||||||||||
Db   361 lkaaqavaltnsvkel 375


RESULT 13
AAB62734
ID   AAB62734 standard; Protein; 392 AA.
XX
AC   AAB62734;
XX
DT   03-APR-2001   (first entry)
XX
PN   WO200078966-A1.
XX
PD   28-DEC-2000.
XX
PF   19-JUN-2000;   2000WO-US16915.
XX
PR   18-JUN-1999;   99US-0140042.
XX
PA   (UYNY ) UNIV NEW YORK STATE RES FOUND.
PA   (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
OS   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
XX
     Borrelia sp chimeric ospC protein SEQ ID NO: 72.
XX
     Chimeric; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
     Dattwyler RJ, Seinost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
PI   2001-050113/06.
DR   WPI; 2001-050113/06.
DR   N-PSDB; AAF29038.
XX
     Compositions of OspC polypeptides from strains of Borrelia which cause
     Lyme disease are used to immunize animals and detect immune responses
     to Lyme disease.
```

RESULT 14

```
PS   Claim 43; Page 139-140; 160pp; English.
XX
CC   The present invention provides compositions comprising ospc proteins and
CC   chimeric ospC proteins from members of the Borrelia genus. These may be
CC   Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC   vaccines against Borrelia infection, which is spread by ticks and leads
CC   to Lyme disease.
XX
SQ   Sequence 392 AA;

Query Match             Best Local Similarity 80.9%;  Score 1490; DB 22; Length 392;
                        Matches 311; Conservative 23; Mismatches 34; Indels 6; Gaps 2;

Qy     2 ACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAI 61
          :|||||||||||||||||||||||||||||||||:|||::||||||||:|||::||:|||
Db    25 scnnsgkdgnasansadesvkgpnlteiskkitesnavvlavkevetllasidelatkai 84

Qy    62 GKIHQNNGLDTEYNHNGSLLAGAYAISTLLIKQKLDGLKNEGLKEKIDAAKKCSETFTNK 121
          ||:| |||||:  |: | ||:||||||:|:|  ||:|||||:||||||||:|||:| |||
Db    85 gkki-gnngleanqsknts-llsgayaisdliaeklnvlkneelkekidtakgcseftnk 143

Qy   122 LKEKHTDLGKEGVTDADAKEAILKTNGTKTKTGAEELGKLFESVEVLSKRAAKEMLANSVK 181
          ||:|| :||:::: |::|:: ||||:|:|:|||||||:| ||:||:|||:||:: |:|||:
Db   144 lksehavlgldnltddnaqrailkkhankdkgaaelekifkavenlskaaqdtlknavke 203

Qy   182 LTSPVVAESPAMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVETLL 241
          |||| ||      |||||||||||||||||||||||||||||||||||||||||||||||
Db   204 ltspivhg-----nnsgkdgntsansadesvkgpnlteiskkitesnavvlavkevetll 258

Qy   242 TSIDELAKAIGKKIKINDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKRAEIEK 301
          |||||||||||||||:|||||||||||||||||||||||||||||||||||||||:|||||
Db   259 tsidelakaigkkikndvsldneadhngslisgaylisnlitkkisaikdsgelkaeiek 318

Qy   302 AKKCSEEFTAKLKGEHTDLGKEGVTDDNAKRAILKTNNDKTKGADELEKLFESVKNLSKA 361
          |||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||
Db   319 akkcseeftaklkgehtdlgkegvtddnakrailktnndktkgadeleklfesvknlska 378

Qy   362 AKEMLTNSVKELTS 375
          ||||||||||||||
Db   379 akemltnsvkelts 392


RESULT 14
AAB62726
ID   AAB62726 standard; Protein; 384 AA.
XX
AC   AAB62726;
XX
DT   03-APR-2001   (first entry)
XX
OS   Borrelia sp chimeric ospc protein SEQ ID NO: 56.
DE
XX
KW   Chimeric; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
XX
PN   WO200078966-A1.
XX
PD   28-DEC-2000.
XX
PR   18-JUN-1999;   99US-0140042.
XX
PF   19-JUN-2000;   2000WO-US16915.
XX
PA   (UYNY ) UNIV NEW YORK STATE RES FOUND.
PA   (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
     Dattwyler RJ, Seinost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
DR   WPI; 2001-050113/06.
```

Left column

DR   N-PSDB; AAF29030.
XX
PT   Compositions of OspC polypeptides from strains of Borrelia which cause
PT   Lyme disease are used to immunize animals and detect immune responses
PT   to Lyme disease -
XX
PS
XX
XX
CC   The present invention provides compositions comprising ospC proteins and
CC   chimeric ospC proteins from members of the Borrelia genus. These may be
CC   Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC   vaccines against Borrelia infection, which is spread by ticks and leads
CC   to Lyme disease.
XX
CC
XX
PS   Claim 43; Page 117-118; 160pp; English.
XX
SQ   Sequence   384 AA;

Query Match   80.8%;   Score 1488;   DB 22;   Length 384;
Best Local Similarity 82.0%;   Pred. No. 4.1e-95;
Matches 309; Conservative 31; Mismatches 31; Indels 6; Gaps 3;

```
Qy    2 ACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAKAI 61
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 acnnsgkdgntsansadesvkgpnlteiskkitdsnavllavkeveallssideiakai 60

Qy   62 GKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNK 121
          |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db   61 gkkihqnngldteynhngsllagayaistllikqkldglkneglkekidaakkcsetftnk 121

Qy  122 LKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 181
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 lkekhtdlgkegvtdadakeailktngtktkgaeelgklfesvevlskaakemlansvke 180

Qy  182 LTSPVVAESPAMV---NNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVE 238
          ||||||||||        ||| |||||||||||||||||||||||||||||||||||||
Db  181 ltspvvaespkkpfhgnnsg--gdsaastnpdesaakgpnltviskkitdsnafllavkeve 238

Qy  239 TLLTSIDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAE 298
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  239 lkeakdcsqkfctkldshaelgigsvgddnakkailkthgtkdkgakeleelfkslesl 297

Qy  299 IEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNL 358
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  298 ikeakdcsqkftktkldshaelgigsvgddnakkailkthgtkdkgakeleelfkslesl 357

Qy  359 SKAAKEMLTNSVKELTS 375
          ||||| ||||||||||:
Db  358 skaaqaaltnsvkeltn 374
```

Right column

PA   (UYNY) UNIV NEW YORK STATE RES FOUND.
PA   (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI   Dattwyler RJ, Seinost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
DR   WPI; 2001-050113/06.
DR   N-PSDB; AAF29041.
XX
PT   Compositions of OspC polypeptides from strains of Borrelia which cause
PT   Lyme disease are used to immunize animals and detect immune responses
PT   to Lyme disease -
XX
PS
XX
CC   The present invention provides compositions comprising ospC proteins and
CC   chimeric ospC proteins from members of the Borrelia genus. These may be
CC   Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC   vaccines against Borrelia infection, which is spread by ticks and leads
CC   to Lyme disease.
XX
CC
XX
PS   Claim 43; Page 147-148; 160pp; English.
XX
SQ   Sequence   408 AA;

Query Match   80.6%;   Score 1485;   DB 22;   Length 408;
Best Local Similarity 81.7%;   Pred. No. 7.2e-95;
Matches 308; Conservative 32; Mismatches 31; Indels 6; Gaps 3;

```
Qy    2 ACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAKAI 61
          :||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   25 scnnsgkdgntsansadesvkgpnlteiskkitdsnavllavkeveallssideiakai 84

Qy   62 GKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNK 121
          |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db   85 gkkihqnngldteynhngsllagayaistllikqkldglkneglkekidaakkcsetftnk 144

Qy  122 LKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 181
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  145 lkekhtdlgkegvtdadakeailktngtktkgaeelgklfesvevlskaakemlansvke 204

Qy  182 LTSPVVAESPAMV---NNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVE 238
          ||||||||||        ||| |||||||||||||||||||||||||||||||||||||
Db  205 ltspvvaespkkpfhgnnsg--gdsaastnpdesaakgpnltviskkitdsnafllavkeve 262

Qy  239 TLLTSIDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAE 298
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  263 allssidelskaigkkikndgtldneanrnesllagayeiskiltqklsvl-nseelkkk 321

Qy  299 IEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNL 358
          :|||||||||||| ||:||:|||  ||||||||||:: |||| |||||||||||||||||:
Db  322 ikeakdcsqkftktkldshaelgigsvgddnakkailkthgtkdkgakeleelfkslesl 381

Qy  359 SKAAKEMLTNSVKELTS 375
          |||| ||||||||||||:
Db  382 skaaqaaltnsvkeltn 398
```

Bottom left

RESULT 15
ID   AAB62737 standard; Protein; 408 AA.
XX
AC   AAB62737;
XX
DT   03-APR-2001 (first entry)
XX
DE   Borrelia sp chimeric ospC protein SEQ ID NO: 78.
XX
KW   Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
XX
PN   WO200078966-A1.
XX
PD   28-DEC-2000.
XX
PF   19-JUN-2000; 2000WO-US16915.
XX
PR   18-JUN-1999; 99US-0140042.
XX

Bottom right

Search completed: March 18, 2002, 09:54:29
Job time: 327 sec

OM protein - protein search, using sw model

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

Run on: March 18, 2002, 09:49:47 ; Search time 55.5 Seconds
(without alignments)
152.049 Million cell updates/sec

Title: US-09-596-746A-24
Perfect score: 1842
Sequence: 1 MACNNSGKDGNTSANSADES..........KNLSKAAKEMLTNSVKELTS 375

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252
Maximum DB seq length: 0
Minimum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :   Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	964	52.3	466	4	US-08-235-836C-110	Sequence 110, App
2	932	50.6	210	4	US-08-235-836C-30	Sequence 30, Appl
3	928	50.4	209	4	US-09-196-293-15	Sequence 15, Appl
4	928	50.4	210	1	US-08-158-353-3	Sequence 3, Appl
5	927	50.3	588	4	US-08-235-836C-122	Sequence 122, App
6	921	50.0	210	4	US-08-209-603E-15	Sequence 15, Appl
7	915	49.7	466	4	US-08-235-836C-107	Sequence 107, App
8	676	36.7	212	4	US-08-158-353-2	Sequence 2, Appl
9	661	35.9	212	4	US-08-031-295-2	Sequence 2, Appl
10	661	35.9	212	4	US-07-903-580-2	Sequence 2, Appl
11	614	33.3	212	4	US-08-158-353-4	Sequence 4, Appl
12	612	33.2	212	4	US-09-196-293-11	Sequence 11, Appl
13	612	33.2	212	4	US-08-209-603E-11	Sequence 11, Appl
14	609	33.1	212	4	US-09-364-083-2	Sequence 2, Appl
15	600.5	32.6	209	4	US-08-600-982-24	Sequence 24, Appl
16	583.5	31.7	207	4	US-08-235-836C-32	Sequence 32, Appl
17	583.5	31.7	209	4	US-08-235-836C-36	Sequence 36, Appl
18	163.5	8.8	2482	1	US-08-328-254-6	Sequence 6, Appli
19	162.5	8.8	3248	2	US-08-353-700-1	Sequence 1, Appli
20	162.5	8.8	3248	4	PCT-US95-16216-1	Sequence 1, Appl
21	157	8.5	194	4	US-09-364-083-2	Sequence 2, Appl
22	148.5	8.1	1713	5	US-08-600-982-24	Sequence 24, Appl
23	148.5	8.1	1713	4	PCT-US94-10261A-24	Sequence 24, Appl
24	147	8.0	1164	5	US-08-923-992A-2	Sequence 2, Appli
25	147	7.9	1196	1	US-08-144-121-4	Sequence 4, Appli
26	146	7.9	1196	2	US-08-733-893-1	Sequence 1, Appli
27	146	7.9	1388	2	US-08-685-576-1	Sequence 1, Appli
28	145	7.9	1588	5	PCT-US93-07261-11	Sequence 11, Appl
29	145	7.9	1663	5	PCT-US93-07261-16	Sequence 16, Appl
30	145	7.9	3111	2	US-08-460-309-4	Sequence 4, Appli
31	143.5	7.8	3111	1	US-08-125-077-4	Sequence 4, Appli
32	143	7.8	630	4	US-08-973-462-9	Sequence 9, Appli
33	142.5	7.7	1164	4	US-08-923-992A-10	Sequence 10, App
34	141	7.7	688	3	US-09-141-047-8	Sequence 8, Appli
35	139	7.5	1786	4	US-08-973-462-8	Sequence 8, App
36	138.5	7.5	2285	4	US-09-308-375-2	Sequence 2, Appli
37	138	7.5	1098	4	US-08-923-992A-8	Sequence 8, Appli
38	137	7.4	1128	4	US-08-923-992A-6	Sequence 6, Appli
39	136	7.4	1104	4	US-08-585-576-4	Sequence 4, Appl
40	136	7.4	1388	2	US-08-235-836C-112	Sequence 112, Appl
41	135.5	7.3	573	4	US-08-961-083-160	Sequence 160, Appl
42	133.5	7.2	641	4	US-08-894-017-23	Sequence 23, Appl
43	132.5	7.2	1561	5	US-09-085-199B-11	Sequence 11, Appl
44	131	7.1	1068	5	PCT-US93-03077-1	Sequence 1, Appli
45	130.5	7.1	396	1	US-08-430-024-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-08-235-836C-110
Sequence 110, Application US/0825836C
Patent No. 6248562

; Patent No. 6248562
; GENERAL INFORMATION:
; APPLICANT: Dunn, John J.
; APPLICANT: Luft, Benjamin J.
; TITLE OF INVENTION: No. 6248562e1 Chimeric Proteins Comprising
; TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor
; NUMBER OF SEQUENCES: 144
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brookhaven National Laboratory
; STREET:
; CITY: Upton
; STATE: NY
; COUNTRY: USA
; ZIP: 11973
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/235,836C
; APPLICATION NUMBER: US/08/235,836C
; FILING DATE: 29-APR-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/148,191
; FILING DATE: 01-11-93
; ATTORNEY/AGENT INFORMATION:
; NAME: Bogosian, Margaret C.
; REGISTRATION NUMBER: 25,324
; REFERENCE/DOCKET NUMBER: BNL93-28A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 282-7338
; TELEFAX: (516) 282-3729
; INFORMATION FOR SEQ ID NO: 110:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 466 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-235-836C-110

Query Match 52.3%; Score 964; DB 4; Length 466;
Best Local Similarity 59.6%; Pred. No. 7e-62;
Matches 229; Conservative 31; Mismatches 92; Indels 32; Gaps 9;

Qy   1 MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA 60

RESULT 2

```
Db   17  ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA   76
Qy   61  IGKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN  120
Db   77  IGKKIHQNNGLDTEYNHINGSLLAGRYAISTLIKQKLDGLKNEGLREKIDAAKKCSETFTN 136
Qy  121  KLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK  180
Db  137  KLKAKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK  196
Qy  181  ELTSPVVAESP--AMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNA--VVLAVKE  236
Db  197  ELTSPVVAESPKKPKQNVSSLD---EKNSVSVDLPGEMKVLVSKEKNKDGKYDLIATVDK  253
Qy  237  VETLLTS------IDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLI--SNLITKKIS   287
Db  254  LELKGTSDKNNGSGVLEGVKADKSKVKLTIS-----DDLGQTTLEVFKEDGKTLVSKKVT  308
Qy  288  AIKDSGELKAEIEKAKKCSEEFTAKLKGEHTDLGKEGYVTDDNAKKAILKTNNDKTKGADE 347
Db  309  S-KDKSSTEEKFNEKGEVSEKIITRADG--TRLEYTGIKSDGSGKAKEVLKGTVLEGTLTI 365
Qy  348  LEKLFESV-------KNLSKAAK                                       363
Db  366  AEKITLLVVKEGTVTLSKNISKSGE                                     389
```

```
RESULT 2
US-08-235-836C-30
Sequence 30, Application US/08235836C
Patent No. 6248562
GENERAL INFORMATION:
    APPLICANT:
        APPLICANT: Dunn, John J.
        APPLICANT: Luft, Benjamin J
    TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor
    TITLE OF INVENTION: No. 62485621el Chimeric Proteins Comprising
    NUMBER OF SEQUENCES: 144
    CORRESPONDENCE ADDRESS:
        ADDRESSEE: Brookhaven National Laboratory
        STREET:
        CITY: Upton
        STATE: NY
        COUNTRY: USA
        ZIP: 11973
    COMPUTER READABLE FORM:
        MEDIUM TYPE: Floppy disk
        COMPUTER: IBM PC compatible
        OPERATING SYSTEM: PC-DOS/MS-DOS
        SOFTWARE: Patentin Release #1.0, Version #1.25
    CURRENT APPLICATION DATA:
        APPLICATION NUMBER: US/08/235,836C
        FILING DATE: 29-APR-1994
    PRIOR APPLICATION DATA:
        APPLICATION NUMBER: 435
        CLASSIFICATION:
        FILING DATE: 01-11-93
        APPLICATION NUMBER: US 08/148,191
    ATTORNEY/AGENT INFORMATION:
        NAME: Bogosian, Margaret C.
        REGISTRATION NUMBER: 25,324
        REFERENCE/DOCKET NUMBER: BNL93-28A
    TELECOMMUNICATION INFORMATION:
        TELEPHONE: (516) 282-7338
        TELEFAX: (516) 282-3729
INFORMATION FOR SEQ ID NO: 30:
    SEQUENCE CHARACTERISTICS:
        LENGTH: 210 amino acids
        TYPE: amino acid
        TOPOLOGY: linear
        MOLECULE TYPE: protein
US-08-235-836C-30
```

RESULT 3

```
Query Match
Best Local Similarity  50.6%;   Score 932;   DB 4;   Length 210;
Matches 188;  Conservative  2;  Mismatches 4.9e-60;  Indels 0;  Gaps 0;

Qy    1  MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA   60
Db   17  ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA   76
Qy   61  IGKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN  120
Db   77  IGKKIHQNNGLDTEYNHINGSLLAGRYAISTLIKQKLDGLKNEGLREKIDAAKKCSETFTN 136
Qy  121  KLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK  180
Db  137  KLKAKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK  196
Qy  181  ELTSPVVAESP                                                   191
Db  197  ELTSPVVAESP                                                   207
```

```
RESULT 3
US-09-196-293-15
Sequence 15, Application US/09196293
Patent No. 6187755
GENERAL INFORMATION:
    APPLICANT: Fuchs, Renate
    APPLICANT: Motz, Manfred
    APPLICANT: Soutscheck, Erwin
    APPLICANT: Wilske, Bettina
    APPLICANT: Preac-Mursic, Vera
    TITLE OF INVENTION: Active proteins from Borrelia
    FILE REFERENCE: 738.001US2
    CURRENT APPLICATION NUMBER: burgdorferi
    CURRENT APPLICATION NUMBER: US/09/196,293
    CURRENT FILING DATE: 1998-11-19
    EARLIER APPLICATION NUMBER: US 08/209,603
    EARLIER FILING DATE: 1994-03-10
    EARLIER APPLICATION NUMBER: US 07/862,535
    EARLIER FILING DATE: 1992-06-19
    EARLIER APPLICATION NUMBER: WO PCT/EP90/02282
    EARLIER FILING DATE: 1990-12-21
    EARLIER APPLICATION NUMBER: DE P39 42 728.5
    EARLIER FILING DATE: 1989-12-22
    EARLIER APPLICATION NUMBER: DE P40 18 988.0
    EARLIER FILING DATE: 1990-06-13
    NUMBER OF SEQ ID NOS: 16
    SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 15
    LENGTH: 209
    TYPE: PRT
    ORGANISM: Borrelia burgdorferi
US-09-196-293-15
```

```
Query Match
Best Local Similarity  50.4%;   Score 928;   DB 4;   Length 209;
98.4%;   Pred. No. 9.5e-60;
Matches 188;  Conservative  2;  Mismatches 1;  Indels 0;  Gaps 0;

Qy    1  MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA   60
Db   17  ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA   76
Qy   61  IGKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN  120
Db   77  IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN  136
Qy  121  KLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK  180
Db  137  KLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK  196
Qy  181  ELTSPVVAESP                                                   191
```

Db   197 ELTSPVVAESP 207

RESULT 4
US-08-158-353-3
Sequence 3, Application US/08158353
Patent No. 5620862
; GENERAL INFORMATION:
;   APPLICANT: Padula, Steven J.
;   TITLE OF INVENTION: Methods for Diagnosing Early Lyme
;   NUMBER OF SEQUENCES: 7
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
;   STREET: Two Militia Drive
;   CITY: Lexington
;   STATE: MA
;   COUNTRY: USA
;   ZIP: 02173
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: Patentin Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/158,353
;   FILING DATE:
;   CLASSIFICATION: 435
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Carroll, Alice O.
;   REGISTRATION NUMBER: 33,542
;   REFERENCE/DOCKET NUMBER: UCT93-05
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 617-861-6240
;   TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 3:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 210 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-08-158-353-3

Query Match   50.4%; Score 928; DB 1; Length 210;
Best Local Similarity   98.4%; Pred. No. 9.5e-60;
Matches 188; Conservative   2; Mismatches   1; Indels   0; Gaps   0;

Qy   1 MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA 60
             :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   17 ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA 76

Qy   61 IGKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN 120
             ||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db   77 IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN 136

Qy   121 KLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   137 KLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 196

Qy   181 ELTSPVVAESP 191
             |||||||||||
Db   197 ELTSPVVAESP 207

RESULT 5
US-08-235-836C-122
Sequence 122, Application US/0823586C
Patent No. 6248562
; GENERAL INFORMATION:
;   APPLICANT: Dunn, John J.
;   APPLICANT: Luft, Benjamin J.

;   TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising
;   TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor
;   NUMBER OF SEQUENCES: 144
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Brookhaven National Laboratory
;   STREET: Upton
;   CITY:
;   STATE: NY
;   COUNTRY: USA
;   ZIP: 11973
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: Patentin Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/235,836C
;   FILING DATE: 29-APR-1994
;   CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 08/148,191
;   FILING DATE: 01-11-93
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Bogoslan, Margaret C.
;   REGISTRATION NUMBER: 25,324
;   REFERENCE/DOCKET NUMBER: BNL93-28A
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (516) 282-7338
;   TELEFAX: (516) 282-3729
; INFORMATION FOR SEQ ID NO: 122:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 588 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-08-235-836C-122

Query Match   50.3%; Score 927; DB 4; Length 588;
Best Local Similarity   88.9%; Pred. No. 4.3e-59;
Matches 193; Conservative   4; Mismatches 10; Indels 10; Gaps   2;

Qy   4 NNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGK 63
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   286 NNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGK 345

Qy   64 KIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNKLK 123
             |||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db   346 KIHQNNGLDTEYNHNGSLLAGRYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNKLK 405

Qy   124 EKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKELT 183
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   406 EKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKELT 465

Qy   184 SPVVAESP-------AMVNN----SGKDGNTSANSADE 210
             |||||||| | | : :||
Db   466 SPVVAESPKPGTMAQYNQHMMLSNKSASQNVRTAEE 502

RESULT 6
US-08-209-603E-15
Sequence 15, Application US/08209603E
Patent No. 6248538
; GENERAL INFORMATION:
;   APPLICANT: FUCHS, RENATE
;   APPLICANT: WILSKE, BETTINA
;   APPLICANT: PREAC-MURSIC, VERA
;   APPLICANT: MOTZ, MANFRED
;   APPLICANT: SOUTSCHECK, ERWIN
;   TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PROTEINS
;   TITLE OF INVENTION: FROM BORRELIA BURGDORFERI
;   NUMBER OF SEQUENCES: 15
;   CORRESPONDENCE ADDRESS:

ADDRESSEE: BROOKS HAIDT HAFFNER & DELAHUNTY
STREET: 99 PARK AVENUE
CITY: NEW YORK
STATE: NY
COUNTRY: USA
ZIP: 10016
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" FLOPPY DISC
COMPUTER: AT&T - IBM COMPATIBLE
OPERATING SYSTEM: MS-DOS Version 6.2
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/209,603E
FILING DATE: 10-MAR-1994
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP90/02282
FILING DATE: 21-DEC-1990
APPLICATION NUMBER: US 07/862,535
FILING DATE: 19-JUN-1992
ATTORNEY/AGENT INFORMATION:
NAME: ROBINSON, WILLIAM R.
REGISTRATION NUMBER: 27,224
REFERENCE/DOCKET NUMBER: LKR-9217-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 697-3355
TELEFAX: (212) 557-5635
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 210
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
DESCRIPTION: PROTEIN
HYPOTHETICAL: N/A
ANTI-SENSE: N/A
FRAGMENT TYPE: N/A
ORIGINAL SOURCE:
ORGANISM: B. BURGDORFERI
IMMEDIATE SOURCE:
LIBRARY: DSM 5662
POSITION IN GENOME: N/A
FEATURE:
IDENTIFICATION METHOD: N/A
PUBLICATION INFORMATION:
US-08-209-603E-15
IDENTIFICATION METHOD: amino acid analysis

Query Match
Best Local Similarity 50.0%; 97.8%;
Score 921; DB 4; Length 210; Pred. No. 3e-59;
Matches 187; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```
Qy    1 MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA 60
          :|||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db   17 ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVLEVEALLSSIDEIAAKA 76

Qy   61 IGKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSEFTN 120
         ||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db   77 IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSEFTN 136

Qy  121 KLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  137 KLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 196

Qy  181 ELTSPVVAESP 191
         |||||||||||
Db  197 ELTSPVVAESP 207
```

RESULT 7
US-08-235-8366-107
; Sequence 107, Application US/08235836C
; Patent No. 6248562

GENERAL INFORMATION:
APPLICANT: Dunn, John J.
APPLICANT: Luft, Benjamin J.
TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising
TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor
NUMBER OF SEQUENCES: 144
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brookhaven National Laboratory
STREET:
CITY: Upton
STATE: NY
COUNTRY: USA
ZIP: 11973
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/235,836C
FILING DATE: 29-APR-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/148,191
FILING DATE: 01-11-93
ATTORNEY/AGENT INFORMATION:
NAME: Bogoslan, Margaret C.
REGISTRATION NUMBER: 25,324
REFERENCE/DOCKET NUMBER: BNL93-28A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 282-3729
TELEFAX: (516) 282-7338
INFORMATION FOR SEQ ID NO: 107:
SEQUENCE CHARACTERISTICS:
LENGTH: 466 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-235-8366-107

Query Match
Best Local Similarity 49.7%; 98.9%;
Score 915; DB 4; Length 466; Pred. No. 2.3e-58;
Matches 186; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy    4 NNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGK 63
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  276 NNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGK 335

Qy   64 KIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSEFTNKLK 123
         |||||||||||||| ||||||| |||||||||||||||||||||||||||||||||||||
Db  336 KIHQNNGLDTEYNINGSLLAGRYAISTLIKQKLDGLKNEGLKEKIDAAKKCSEFTNKLK 395

Qy  124 EKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKELT 183
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  396 AKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKELT 455

Qy  184 SPVVAESP 191
         ||||||||
Db  456 SPVVAESP 463
```

RESULT 8
US-08-158-353-2
; Sequence 2, Application US/08158353
; Patent No. 5620862
GENERAL INFORMATION:
APPLICANT: Padula, Steven J.
TITLE OF INVENTION: Methods for Diagnosing Early Lyme
TITLE OF INVENTION: Disease
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.

STREET: Two Militia Drive
CITY: Lexington
STATE: MA
COUNTRY: USA
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/158,353
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Carroll, Alice O.
REGISTRATION NUMBER: 33,542
REFERENCE/DOCKET NUMBER: UCT93-05
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-861-6240
TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 212 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-158-353-2,

Query Match   36.7%;   Score 676;   DB 1;   Length 212;
Best Local Similarity   74.7%;   Pred. No. 1.2e-41;
Matches 145;   Conservative   18;   Mismatches   27;   Indels   4;   Gaps   3;

Qy   1 MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA 60
          ::|||||||||||||||||||||||||||||||:||||:||||:||||||:|||::|||:||
Db   17 ISCNNSGKDGNTSANSADESVKGPNLTEISKKITPESNAVVLAVKEVETLLASIDEVAKRA 76

Qy   61 IGKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKN-EGLKEKIIDAAKKCSETFT 119
          |:| |||:|||| |||||||||||||||||||:|||||||::|||||||:|||||||||||||
Db   77 IGNLIAQ-NGLNAGANQNGSLLAGAYVISTLIPAEKLDGLKNSEELKEKIEDAKKCNKAFT 135

Qy   120 NKLKEKHTDLG--KEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLAN 177
          :||| | :|| |||:|| |:||||||:|||||||||||||||||:|||||:|| | |
Db   136 DKLKSSHAELGIANGAATDANAKAAILKTNGTKDKGAQELEKLFESVKNLSKAAQETLNN 195

Qy   178 SVKELTSPVVAESP 191
          ||||||||||||:|
Db   196 SVKELTSPVVAENP 209

RESULT 9
US-08-031-295-2
Sequence 2, Application US/08031295
Patent No. 5530103
GENERAL INFORMATION:
APPLICANT: LIVEY, Ian
APPLICANT: DORNER, Freidrich
TITLE OF INVENTION: METHOD AND COMPOSITION FOR THE
TITLE OF INVENTION: PREVENTION OF LYME DISEASE
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 2007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/031,295
FILING DATE: 19930312
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/903,580
FILING DATE: 25-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/824,161
FILING DATE: 22-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/727,245
FILING DATE: 11-JUL-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/142 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 212 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-031-295-2

Query Match   35.9%;   Score 661;   DB 1;   Length 212;
Best Local Similarity   73.6%;   Pred. No. 1.5e-40;
Matches 142;   Conservative   14;   Mismatches   35;   Indels   2;   Gaps   2;

Qy   1 MACNNSGKDGNT-SANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAK 59
          ::|||||| |:: | ||| ||||||||||||||||||||||:||||||||| :::||:|:|
Db   17 ISCNNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETLVSSIDELATK 76

Qy   60 AIGKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEG-LKEKIDAAKKCSETF 118
          ||||||:|||| |||: |||||||||||||||||||||||||||| ||||:|||||||||
Db   77 AIGKKIQQNNGLGANADKNGSLLAGAYAISTLITEKLKALKNSGELKAKIEDAKKCSEDF 136

Qy   119 TNKLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANS 178
          |:||||:|||||||||| | |||:|||||||:|||||||||||||:|||||||:|| | |
Db   137 TKKLAAGHAQLGIDGATDNDSKEAILKTNGTKTKGARELVKLSESVASLSKAAQEASANS 196

Qy   179 VKELTSPVVAESP 191
          ||||||||||:|
Db   197 VKELTSPVVAETP 209

RESULT 10
US-07-903-580-2
Sequence 2, Application US/07903580
Patent No. 6221363
GENERAL INFORMATION:
APPLICANT: LIVEY, Ian
APPLICANT: DORNER, Freidrich
TITLE OF INVENTION: METHOD AND COMPOSITION FOR THE
TITLE OF INVENTION: PREVENTION OF LYME DISEASE
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: Virginia
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/07/903,580
  FILING DATE: 19920625
  CLASSIFICATION: 435
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 07/824,161
  FILING DATE: 22-JAN-1992
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 07/727,245
  FILING DATE: 11-JUL-1991
ATTORNEY/AGENT INFORMATION:
  NAME: BENT, Stephen A.
  REGISTRATION NUMBER: 29,768
  REFERENCE/DOCKET NUMBER: 30472/131 IMMU
TELECOMMUNICATION INFORMATION:
  TELEPHONE: (703)836-9300
  TELEFAX: (703)683-4109
  TELEX: 899149
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
  LENGTH: 212 amino acids
  TYPE: AMINO ACID
  TOPOLOGY: linear
  MOLECULE TYPE: protein
US-07-903-580-2

Query Match   35.9%; Score 661; DB 4; Length 212;
Best Local Similarity   73.6%; Pred. No. 1.5e-40;
Matches 142; Conservative 14; Mismatches 35; Indels 2; Gaps 2;

```
Qy     1 MACNNSGKDGNT-SANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAK 59
           :||||:|| ||::| :|| |-||||||||||||||||||||||||||||||:|||||:| |
Db    17 ISCNNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETLVSSIDELATK 76

Qy    60 AIGKKIHQNNGLDTEYNHNGSLLAGAYAISTLLIKQKLDGLKNEG-LKEKIDAAKKCSETF 118
          |||:|| ||||  : |||||||||||||||||:|||| ||:|| ||:||||||||||| |
Db    77 AIGKKIQQNNGLGANADKNGSLLAGAYAISTLLITEKLKALKNSGELKAKIEDAKKCSEDF 136

Qy   119 TNKLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANS 178
          ||| | |||| |||||| ||||||||||||||||||||||||||||||||||| ||
Db   137 TKKLAAGHAQLGIDGATDNDSKEAILKTNGTKTKGAEELVKLSESVASLSKRAQEASANS 196

Qy   179 VKELTSPVVAESP 191
          |||||||||||:|
Db   197 VKELTSPVVAETP 209
```

RESULT 11
US-08-158-353-4
Sequence 4, Application US/08158353
Patent No. 5620862
GENERAL INFORMATION:
APPLICANT: Padula, Steven J.
TITLE OF INVENTION: Methods for Diagnosing Early Lyme
TITLE OF INVENTION: Disease
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
  ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
  STREET: Two Militia Drive
  CITY: Lexington
  STATE: MA
  COUNTRY: USA
  ZIP: 02173
COMPUTER READABLE FORM:
  MEDIUM TYPE: Floppy disk
  COMPUTER: IBM PC compatible
  OPERATING SYSTEM: PC-DOS/MS-DOS
  SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/158,353
  FILING DATE:
  CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
  NAME: Carroll, Alice O.
  REGISTRATION NUMBER: 33,542
  REFERENCE/DOCKET NUMBER: UCT93-05
TELECOMMUNICATION INFORMATION:
  TELEPHONE: 617-861-6240
  TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
  LENGTH: 212 amino acids
  TYPE: amino acid
  STRANDEDNESS: single
  TOPOLOGY: linear
  MOLECULE TYPE: protein
US-08-158-353-4

Query Match   33.3%; Score 614; DB 1; Length 212;
Best Local Similarity   69.9%; Pred. No. 3.5e-37;
Matches 135; Conservative 14; Mismatches 42; Indels 2; Gaps 2;

```
Qy     1 MACNNSGKDGNT-SANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAK 59
           ::|||:|| : | ||| |-|||||||||||||||||||||||||||:|||||:: |
Db    17 ISCNNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETLVSSIDELAKK 76

Qy    60 AIGKKIHQNNGLDTEYNHNGSLLAGAYAISTLLIKQKLDGLKN-EGLKEKIDAAKKCSETF 118
          |||:|| |||| |||||||||||||||||:|||| ||| :|:|||||||||||| |
Db    77 AIGQKIDNNNGLAALNNQNGSLLAGAYAISTLITEKLSKLKNLELKTEIAKAKKCSEEF 136

Qy   119 TNKLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANS 178
          ||| | |||| |:| |: | ||||||||||||||||||:| ||:| :||||| ||
Db   137 TNKLKSGHADLGKQDATDDHAKAAILKTHAFTDKGAKEFKDLFESVEGLIKAAQVALTNS 196

Qy   179 VKELTSPVVAESP 191
          |||||||||||||
Db   197 VKELTSPVVAESP 209
```

RESULT 12
US-09-196-293-11
Sequence 11, Application US/09196293
Patent No. 6183755
GENERAL INFORMATION:
APPLICANT: Fuchs, Renate
APPLICANT: Motz, Manfred
APPLICANT: Soutscheck, Erwin
APPLICANT: Wilske, Bettina
APPLICANT: Preac-Mursic, Vera
TITLE OF INVENTION: Active proteins from Borrelia
TITLE OF INVENTION: burgdorferi
FILE REFERENCE: 738.001US2
GENERAL INFORMATION:
CURRENT APPLICATION DATA:
CURRENT FILING DATE: 1998-11-19
CURRENT APPLICATION NUMBER: US/09/196,293
EARLIER APPLICATION NUMBER: US 08/209,603
EARLIER FILING DATE: 1994-03-10
EARLIER APPLICATION NUMBER: US 07/862,535
EARLIER FILING DATE: 1992-06-19
EARLIER APPLICATION NUMBER: WO PCT/EP90/02282
EARLIER FILING DATE: 1990-12-21
EARLIER APPLICATION NUMBER: DE P39 42 728.5
EARLIER FILING DATE: 1989-12-22
EARLIER APPLICATION NUMBER: DE P40 18 988.0
EARLIER FILING DATE: 1990-06-13
NUMBER OF SEQ ID NOS: 16
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 11
  LENGTH: 212
  TYPE: PRT
  ORGANISM: Borrelia burgdorferi
US-09-196-293-11

Query Match   33.2%; Score 612; DB 4; Length 212;

RESULT 13

```
Best Local Similarity  70.5%;  Pred. No. 4.8e-37;  Mismatches  43;  Indels  2;  Gaps  2;
Matches 136;  Conservative 12;
```

```
Qy    1  MACNNSGKDG-NTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAK  59
          ::|||||||| | |||||||||||:||||||||||||||||||||||:||||||::|||||||
Db   17  ISCNNSGKVGILTSTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETLVLSIDELAKK  76
```

```
Qy   60  AIGKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKN-EGLKEKIDAAKKCSETF  118
          |||:||| |||| | ||||||||||||||||||||:|||||| ||||||:|| |||||||
Db   77  AIGOKIDNNNGLAALNNONGSLLAGAYAISTLITEKLSKLKNLEELKTEIAKAKKCSEEF  136
```

```
Qy  119  TNKLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANS  178
          |||| ||||||||||:|| |||||||||||||||||||||||||||:||| ||| ||
Db  137  TNKLKSGHADLGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGLLKAAQVALTNS  196
```

```
Qy  179  VKELTSPVVAESP  191
          |||||||||||||
Db  197  VKELTSPVVAESP  209
```

```
Qy ...
```

```
GENERAL INFORMATION:
Patent No. 6248538
Sequence 11, Application US/08209603E
CORRESPONDENCE ADDRESS:
    ADDRESSEE:  BROOKS HAIDT HAFFNER & DELAHUNTY
    STREET:  99 PARK AVENUE
    CITY:  NEW YORK
    STATE:  NY
    COUNTRY:  USA
    ZIP:  10016
COMPUTER READABLE FORM:
    MEDIUM TYPE:  3.5" FLOPPY DISC
    COMPUTER:  AT&T - IBM COMPATIBLE
    OPERATING SYSTEM:  MS-DOS Version 6.2
    SOFTWARE:  ASCII
CURRENT APPLICATION DATA:
    APPLICATION NUMBER:  US/08/209,603E
    FILING DATE:  10-MAR-1994
    CLASSIFICATION:  436
PRIOR APPLICATION DATA:
    APPLICATION NUMBER:  PCT/EP90/02282
    FILING DATE:  21-DEC-1990
    APPLICATION NUMBER:  US 07/862,535
    FILING DATE:  19-JUN-1992
ATTORNEY/AGENT INFORMATION:
    NAME:  ROBINSON, WILLIAM R.
    REGISTRATION NUMBER:  27,224
    REFERENCE/DOCKET NUMBER:  LKR-9217-A
TELECOMMUNICATION INFORMATION:
    TELEPHONE:  (212) 697-3355
    TELEFAX:  (212) 557-5635
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
    LENGTH:  212
    TYPE:  AMINO ACID
    TOPOLOGY:  LINEAR
MOLECULE TYPE:
    DESCRIPTION:  PROTEIN
HYPOTHETICAL:  N/A
ANTI-SENSE:  N/A
FRAGMENT TYPE:  N/A
ORIGINAL SOURCE:
    ORGANISM:  B. BURGDORFERI
```

```
APPLICANT:  FUCHS, RENATE
APPLICANT:  WILSKE, BETTINA
APPLICANT:  PREAC-MURSIC, VERA
APPLICANT:  MOTZ, MANFRED
APPLICANT:  SOUTSCHEK, ERWIN
TITLE OF INVENTION:  IMMUNOLOGICALLY ACTIVE PROTEINS
TITLE OF INVENTION:  FROM BORRELIA BURGDORFERI
NUMBER OF SEQUENCES:  15
```

```
IMMEDIATE SOURCE:
    LIBRARY:  DSM 5662
POSITION IN GENOME:  DSM 5662
FEATURE:
    IDENTIFICATION METHOD:  amino acid analysis
    PUBLICATION INFORMATION:  N/A
ANTI-SENSE:  N/A
FRAGMENT TYPE:  INTERNAL
ORIGINAL SOURCE:
US-08-209-603E-11
```

RESULT 14

```
US-08-235-836C-34
Sequence 34, Application US/08235836C
Patent No. 6248562
GENERAL INFORMATION:
APPLICANT:  Dunn, John J.
APPLICANT:  Luft, Benjamin J.
TITLE OF INVENTION:  No. 6248562el Chimeric Proteins Comprising
TITLE OF INVENTION:  Borrelia Polypeptides and Uses Therefor
NUMBER OF SEQUENCES:  144
CORRESPONDENCE ADDRESS:
    ADDRESSEE:  Brookhaven National Laboratory
    STREET:
    CITY:  Upton
    STATE:  NY
    COUNTRY:  USA
    ZIP:  11973
COMPUTER READABLE FORM:
    MEDIUM TYPE:  Floppy disk
    COMPUTER:  IBM PC compatible
    OPERATING SYSTEM:  PC-DOS/MS-DOS
    SOFTWARE:  PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
    APPLICATION NUMBER:  US/08/235,836C
    FILING DATE:  29-APR-1994
    CLASSIFICATION:  435
PRIOR APPLICATION DATA:
    APPLICATION NUMBER:  US 08/148,191
    FILING DATE:  01-11-93
ATTORNEY/AGENT INFORMATION:
    NAME:  Bogoslan, Margaret C.
    REGISTRATION NUMBER:  25,324
    REFERENCE/DOCKET NUMBER:  BNL93-28A
TELECOMMUNICATION INFORMATION:
    TELEPHONE:  (516) 282-7338
    TELEFAX:  (516) 282-3729
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
    LENGTH:  212 amino acids
    TYPE:  amino acid
```

```
Query Match            33.2%;  Score 612;  DB 4;  Length 212;
Best Local Similarity  70.5%;  Pred. No. 4.8e-37;  Mismatches  43;  Indels  2;  Gaps  2;
Matches 136;  Conservative 12;
```

```
Qy    1  MACNNSGKDG-NTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAK  59
          ::|||||||| | |||||||||||:||||||||||||||||||||||:||||||::|||||||
Db   17  ISCNNSGKVGILTSTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETLVLSIDELAKK  76
```

```
Qy   60  AIGKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKN-EGLKEKIDAAKKCSETF  118
          |||:||| |||| | ||||||||||||||||||||:|||||| ||||||:|| |||||||
Db   77  AIGOKIDNNNGLAALNNONGSLLAGAYAISTLITEKLSKLKNLEELKTEIAKAKKCSEEF  136
```

```
Qy  119  TNKLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANS  178
          |||| ||||||||||:|| |||||||||||||||||||||||||||:||| ||| ||
Db  137  TNKLKSGHADLGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGLLKAAQVALTNS  196
```

```
Qy  179  VKELTSPVVAESP  191
          |||||||||||||
Db  197  VKELTSPVVAESP  209
```

;
US-08-235-836C-34
MOLECULE TYPE: protein
TOPOLOGY: linear

Query Match
Best Local Similarity   33.1%;   Score 609; DB 4; Length 212;
Matches 134; Conservative 15; Mismatches 42; Indels 2; Gaps 2;
                         69.4%;   Pred. No. 7.9e-37;

```
Qy    1 MACNNSGKDGNT-SANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAK  59
          ::||:|||||  |::: ||| |||||||||||||||||||||||||||||||::||| |
Db   17 ISCSNSNGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVEFLVLSIDELAKK 76

Qy   60 AIGKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKN-EGLKEKIDAARKKCSETF 118
          |||  ||||||||||||||||||||||||||||||||:|:|| ||||||||||||||| |
Db   77 AIGQKIDNNNGLAALNNONGSLLAGAYAISTLITEKLSKLKNLEELKTEIAKAKKCSEEF  136

Qy  119 TNKLKEKHTDLGKEGVTDADAKEAILKTINGTKTKGAEELGKLFESVEVLSKAAKEMLANS 178
          |||| |||||:|||||||||||||||||||||||||||||||||||||| ||||:||| |
Db  137 TNKLKSGHADLGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGLLKAAQVALTNS  196

Qy  179 VKELTSPVVAESP 191
          |||||||||||||
Db  197 VKELTSPVVAESP 209
```

Matches 131; Conservative 19; Mismatches 39; Indels 3; Gaps 2;

```
Qy    1 MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAK  60
          ::|:|||| |:|:: ||| |||||||||||||:||||||||||||||||||::|||:|| |
Db   17 ISCNNSG--GDTASTNPDESAKGPNLTVISKKITDSNAFVLAVKEVEFLVISIDELANKA  74

Qy   61 IGKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKN-EGLKEKIDAARKKCSETFT 119
          |||||||||||||||||||||||||||||||:||:|| |:|| ||||||||||||||:| |
Db   75 IGKVIHQNNLNANAGQNGSLLAGAYAISTLITEKLSKLKNSEELNKIEEAKNHSEAFT   134

Qy  120 NKLKEKHTDLGKEGVTDADAKEAILKTINGTKTKGAEELGKLFESVEVLSKAAKEMLANSV 179
          :||| ||||||:||||||||||||||||||||||||||||||||||| :||||||||| |
Db  135 NRLKGSHAQLGVAAATDDHAKEAILKSNPTKDKGAKALKDLSESVESLAKAAQEALANSV  194

Qy  180 KELTSPVVAESP 191
          ||||:|:|||||
Db  195 KELTNPVVAESP 206
```

Search completed: March 18, 2002, 09:55:32
Job time: 345 sec

RESULT 15
US-08-235-836C-32
Sequence 32, Application US/0823583GC
Patent No. 6248562
GENERAL INFORMATION:
APPLICANT: Dunn, John J.
APPLICANT: Luft, Benjamin J.
TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising
TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor
NUMBER OF SEQUENCES: 144
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brookhaven National Laboratory
STREET:
CITY: Upton
STATE: NY
COUNTRY: USA
ZIP: 11973
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US 08/235,836C
FILING DATE: 29-APR-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/148,191
FILING DATE: 01-11-93
ATTORNEY/AGENT INFORMATION:
NAME: Bogosian, Margaret C.
REGISTRATION NUMBER: 25,324
REFERENCE/DOCKET NUMBER: BNL93-28A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 282-7338
TELEFAX: (516) 282-3729
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 209 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-235-836C-32

Query Match
Best Local Similarity   32.6%;   Score 600.5; DB 4; Length 209;
                         68.2%;   Pred. No. 3.2e-36;

THIS PAGE BLANK

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 09:52:32 ; Search time 621.2 Seconds
                                           (without alignments)
                                           167.613 Million cell updates/sec

Title: US-09-596-746a-24
Perfect score: 1842
Sequence: 1 MACNNSGKDGNTSANSADES...........KNLSKAAKEMLTNSVKELTS 375

Scoring table: BLOSUM62
               Gapop 10.0 , Gapext 0.5
Searched: 3148936 seqs, 27765703A residues

Total number of hits satisfying chosen parameters: 3148936

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Database :   Pending_Patents_AA_Main:*
    1: /cgn2_6/ptodata/2/paa/PCTUS_COMB.pep:*
    2: /cgn2_6/ptodata/2/paa/US06_COMB.pep:*
    3: /cgn2_6/ptodata/2/paa/US07_COMB.pep:*
    4: /cgn2_6/ptodata/2/paa/US080_COMB.pep:*
    5: /cgn2_6/ptodata/2/paa/US081_COMB.pep:*
    6: /cgn2_6/ptodata/2/paa/US082_COMB.pep:*
    7: /cgn2_6/ptodata/2/paa/US083_COMB.pep:*
    8: /cgn2_6/ptodata/2/paa/US084_COMB.pep:*
    9: /cgn2_6/ptodata/2/paa/US085_COMB.pep:*
   10: /cgn2_6/ptodata/2/paa/US086_COMB.pep:*
   11: /cgn2_6/ptodata/2/paa/US087_COMB.pep:*
   12: /cgn2_6/ptodata/2/paa/US088_COMB.pep:*
   13: /cgn2_6/ptodata/2/paa/US089_COMB.pep:*
   14: /cgn2_6/ptodata/2/paa/US090_COMB.pep:*
   15: /cgn2_6/ptodata/2/paa/US091_COMB.pep:*
   16: /cgn2_6/ptodata/2/paa/US092_COMB.pep:*
   17: /cgn2_6/ptodata/2/paa/US093_COMB.pep:*
   18: /cgn2_6/ptodata/2/paa/US094_COMB.pep:*
   19: /cgn2_6/ptodata/2/paa/US095_COMB.pep:*
   20: /cgn2_6/ptodata/2/paa/US096_COMB.pep:*
   21: /cgn2_6/ptodata/2/paa/US097_COMB.pep:*
   22: /cgn2_6/ptodata/2/paa/US098_COMB.pep:*
   23: /cgn2_6/ptodata/2/paa/US099_COMB.pep:*
   24: /cgn2_6/ptodata/2/paa/US60_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1842	100.0	375	19	US-09-596-746a-24	Sequence 24, Appl
2	1837	99.7	374	19	US-09-596-746-24	Sequence 24, Appl
3	1834	99.6	398	19	US-09-596-746-60	Sequence 60, Appl
4	1834	99.6	399	19	US-09-596-746A-60	Sequence 60, Appl
5	1819.5	98.8	401	19	US-09-596-746-80	Sequence 80, Appl
6	1819.5	98.8	402	19	US-09-596-746-80	Sequence 80, Appl
7	1810.5	98.3	378	19	US-09-596-746A-80	Sequence 80, Appl
8	1810.5	98.3	378	19	US-09-596-746A-28	Sequence 28, Appl
9	1597.5	86.7	374	19	US-09-596-746A-26	Sequence 26, Appl
10	1592.5	86.5	373	19	US-09-596-746-26	Sequence 26, Appl
11	1589.5	86.3	397	19	US-09-596-746A-62	Sequence 62, Appl
12	1589.5	86.3	398	19	US-09-596-746-62	Sequence 62, Appl
13	1575	85.5	400	19	US-09-596-746A-82	Sequence 82, Appl
14	1575	85.5	401	19	US-09-596-746-82	Sequence 82, Appl
15	1566	85.0	377	19	US-09-596-746-30	Sequence 30, Appl
16	1566	85.0	377	19	US-09-596-746A-30	Sequence 30, Appl
17	1534	83.3	371	19	US-09-596-746A-34	Sequence 34, Appl
18	1529	83.0	370	19	US-09-596-746-34	Sequence 34, Appl
19	1526	82.8	394	19	US-09-596-746-64	Sequence 64, Appl
20	1526	82.8	395	19	US-09-596-746A-64	Sequence 64, Appl
21	1498	81.3	369	19	US-09-596-746A-40	Sequence 40, Appl
22	1493	81.1	385	19	US-09-596-746-40	Sequence 40, Appl
23	1493	81.1	393	19	US-09-596-746A-56	Sequence 56, Appl
24	1492	81.0	560	19	US-09-596-746A-52	Sequence 52, Appl
25	1492	81.0	560	19	US-09-596-746-52	Sequence 52, Appl
26	1490	80.9	393	19	US-09-596-746-72	Sequence 72, Appl
27	1490	80.9	393	19	US-09-596-746A-72	Sequence 72, Appl
28	1488	80.8	384	19	US-09-596-746A-56	Sequence 56, Appl
29	1485	80.6	408	19	US-09-596-746-78	Sequence 78, Appl
30	1485	80.6	409	19	US-09-596-746A-78	Sequence 78, Appl
31	1289.5	70.0	370	19	US-09-596-746A-36	Sequence 36, Appl
32	1284.5	69.7	369	19	US-09-596-746-36	Sequence 36, Appl
33	1281.5	69.6	393	19	US-09-596-746-66	Sequence 66, Appl
34	1281.5	69.6	394	19	US-09-596-746A-66	Sequence 66, Appl
35	1281.5	69.6	368	19	US-09-596-746-42	Sequence 42, Appl
36	1253.5	68.1	367	19	US-09-596-746A-42	Sequence 42, Appl
37	1248.5	67.8	391	19	US-09-596-746-74	Sequence 74, Appl
38	1245.5	67.6	392	19	US-09-596-746A-74	Sequence 74, Appl
39	1245.5	67.6	370	19	US-09-596-746-32	Sequence 32, Appl
40	1217.5	66.1	378	19	US-09-596-746A-54	Sequence 54, Appl
41	1213.5	65.9	378	19	US-09-596-746-54	Sequence 54, Appl
42	1213.5	65.9	369	19	US-09-596-746A-32	Sequence 32, Appl
43	1212.5	65.8	369	19	US-09-596-746-68	Sequence 68, Appl
44	1209.5	65.7	393	19	US-09-596-746A-68	Sequence 68, Appl
45	1205.5	65.4	401	19	US-09-596-746-70	Sequence 70, Appl

ALIGNMENTS

RESULT 1
US-09-596-746A-24

Sequence 24, Application US/0959674GA

GENERAL INFORMATION:
APPLICANT: Dattwyler, Raymond J.
APPLICANT: Seinost, Gerald
APPLICANT: Dykhuizen, Daniel
APPLICANT: Luft, Benjamin J.
APPLICANT: Maria J.C. Gomes-Solecki
TITLE OF INVENTION: Groups of Borrelia burgdorferi and
TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
FILE REFERENCE: 2631.1002-001
CURRENT APPLICATION NUMBER: US/09/596,746A
CURRENT FILING DATE: 2000-06-19
PRIOR APPLICATION NUMBER: US 60/140,042
PRIOR FILING DATE: 1999-06-18
NUMBER OF SEQ ID NOS: 84
SOFTWARE: FastSEQ for Windows Version 4.0

SEQ ID NO 24
  LENGTH: 375
  TYPE: PRT
  ORGANISM: Artificial Sequence
  FEATURE:
  OTHER INFORMATION: OspC Chimera
US-09-596-746A-24

Query Match   100.0%; Score 1842; DB 19; Length 375;
Best Local Similarity 100.0%; Pred. No. 8.8e-125;
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 MACNNSGKDGNTSANSADESVKGPNLTEISRKITDSNAVLLAVKEVEALLSSIDEIAAKA 60

RESULT 2
US-09-596-746-24
Sequence 24, Application US/09596746
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seinost, Gerald
; APPLICANT: Dykhuizen, Danial
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: Borrelia afzelii That Cause Lyme Disease in Humans
; CURRENT APPLICATION NUMBER: US/09/596,746
; CURRENT FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: 2631.1002-001
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: OspC Chimera
US-09-596-746-24

Query Match 99.7%; Score 1837; DB 19; Length 374;
Best Local Similarity 100.0%; Pred. No. 2e-124;
Matches 374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Db    1 MACNNSGKDGNTSANSADESVKGPNLTEISKKIITDSNAVLLAVKEVEALLSSIDEIAAKA 60
Qy      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 IGKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN 120
Qy    1 IGKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN 60
Db  121 KLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 180
Qy   61 GKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNK 121
Db  181 ELTSPVVAESPAMVNNSGKDGNTSANSADESVKGPNLTEISKKIITESNAVVLAVKEVETL 240
Qy   62 GKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNK 121
Db  241 LTSIDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAEIE 300
Qy  122 LKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 181
Db  301 KAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLSK 360
Qy  182 LTSPVVAESPAMVNNSGKDGNTSANSADESVKGPNLTEISKKIITESNAVVLAVKEVETLL 241
Db  361 AAKEMLTNSVKELTS 375
Qy      ||||||||||||
```

RESULT 3
US-09-596-746-60
Sequence 60, Application US/09596746
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seinost, Gerald
; APPLICANT: Dykhuizen, Danial
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; FILE REFERENCE: Borrelia afzelii That Cause Lyme Disease in Humans
; CURRENT APPLICATION NUMBER: US/09/596,746
; CURRENT FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: 2631.1002-001
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 60
; LENGTH: 398
; TYPE: PRT
; ORGANISM: OspC Chimera
US-09-596-746-60

Query Match 99.6%; Score 1834; DB 19; Length 398;
Best Local Similarity 99.7%; Pred. No. 3.6e-124;
Matches 373; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Db  181 LTSPVVAESPAMVNNSGKDGNTSANSADESVKGPNLTEISKKIITESNAVVLAVKEVETLL 240
Qy    2 ACNNSGKDGNTSANSADESVKGPNLTEISKKIITDSNAVLLAVKEVEALLSSIDEIAAKA 61
Db  241 TSIDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAEIEK 300
Qy   25 SCNNSGKDGNTSANSADESVKGPNLTEISKKIITDSNAVLLAVKEVEALLSSIDEIAAKAI 61
Db  301 AKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLSKA 361
Qy   62 GKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNK 121
Db  361 AAKEMLTNSVKELTS 374
Qy   85 GKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNK 144
Db  122 LKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 181
Qy  145 LKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 181
Db  182 LTSPVVAESPAMVNNSGKDGNTSANSADESVKGPNLTEISKKIITESNAVVLAVKEVETLL 241
Qy  205 LTSPVVAESPAMVNNSGKDGNTSANSADESVKGPNLTEISKKIITESNAVVLAVKEVETLL 264
Db  242 TSIDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAEIEK 301
Qy  265 TSIDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAEIEK 324
Db  302 AKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLSKA 361
Qy  325 AKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLSKA 384
Db  362 AKEMLTNSVKELTS 375
Qy  385 AKEMLTNSVKELTS 398
```

RESULT 4
US-09-596-746A-60
Sequence 60, Application US/0959674GA
GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Selnost, Gerald
; APPLICANT: Dykhuizen, Danial
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746A
; CURRENT FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 60
; LENGTH: 399
; TYPE: PRT
; ORGANISM: ospC Chimera
US-09-596-746A-60

Query Match   99.6%; Score 1834; DB 19; Length 399;
Best Local Similarity   99.7%; Pred. No. 3. 6e-124;
Matches 373; Conservative   1; Mismatches   0; Indels   0; Gaps   0;

Qy   2 ACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAI 61
           :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   26 SCNNSGKDGNTSANSADESVGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAI 85

Qy   62 GKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNK 121
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   86 GKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNK 145

Qy   122 LKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 181
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   146 LKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 205

Qy   182 LTSPVVAESPAMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVETLL 241
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   206 LTSPVVAESPAMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVETLL 265

Qy   242 TSIDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAEIEK 301
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   266 TSIDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAEIEK 325

Qy   302 AKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLSKA 361
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   326 AKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLSKA 385

Qy   362 AKEMLTNSVKELTS 375
           ||||||||||||||
Db   386 AKEMLTNSVKELTS 399

RESULT 5
US-09-596-746-80
Sequence 80, Application US/0959674GA
GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Selnost, Gerald
; APPLICANT: Dykhuizen, Danial
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746
; CURRENT FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/140,042

; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 80
; LENGTH: 401
; TYPE: PRT
; ORGANISM: ospC Chimera
US-09-596-746-80

Query Match   98.8%; Score 1819; DB 19; Length 401;
Best Local Similarity   98.7%; Pred. No. 2; Mismatches   0; Indels   3; Gaps   1;
Matches 372; Conservative   2;

Qy   2 ACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAI 61
           :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   25 SCNNSGKDGNTSANSADESVGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAI 84

Qy   62 GKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNK 121
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   85 GKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNK 144

Qy   122 LKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 181
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   145 LKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 204

Qy   182 LTSPVVAES---PAMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVE 238
           |||||||| :|||||||||||||||||||||||||||||||||||||||||||||||
Db   205 LTSPVVAESPKKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVE 264

Qy   239 TLLTSIDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAE 298
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   265 TLLTSIDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAE 324

Qy   299 IEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNL 358
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   325 IEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNL 384

Qy   359 SKAAKEMLTNSVKELTS 375
           |||||||||||||||||
Db   385 SKAAKEMLTNSVKELTS 401

RESULT 6
US-09-596-746A-80
Sequence 80, Application US/0959674GA
GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Selnost, Gerald
; APPLICANT: Dykhuizen, Danial
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746A
; CURRENT FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 80
; LENGTH: 402
; TYPE: PRT
; ORGANISM: ospC Chimera
US-09-596-746A-80

Query Match   98.8%; Score 1819.5; DB 19; Length 402;
Best Local Similarity   98.7%; Pred. No. 4.1e-123;
Matches 372; Conservative   2; Mismatches   0; Indels   3; Gaps   1;

Qy   2 ACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAI 61

```
Db   26  SCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKI  85
Qy   62  GKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFINK 121
Db   86  GKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFINK 145
Qy  122  LKEKHTDLGKEGYTDADAKEAILKTNGRTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 181
Db  146  LKEKHTDLGKEGYTDADAKEAILKTNGRTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 205
Qy  182  LTSPVVAES---PAMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVE  238
Db  206  LTSPVVAESPKKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVE  265
Qy  239  TLLTSIDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAE  298
Db  266  TLLTSIDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAE  325
Qy  299  IEKAKKCSEEFTAKLKGEHTDLGKEGYTDDNAKKAILKTNNDKTKGADELEKLFESVKNL  358
Db  326  IEKAKKCSEEFTAKLKGEHTDLGKEGYTDDNAKKAILKTNNDKTKGADELEKLFESVKNL  385
Qy  359  SKAAKEMLTNSVKELTS 375
Db  386  SKAAKEMLTNSVKELTS 402

RESULT 7
US-09-596-746-28
Sequence 28, Application US/0959674G
GENERAL INFORMATION:
  APPLICANT: Dattwyler, Raymond J.
  APPLICANT: Seinost, Gerald
  APPLICANT: Dykhuizen, Daniel
  APPLICANT: Luft, Benjamin J.
  APPLICANT: Maria J.C. Gomes-Solecki
  TITLE OF INVENTION: Groups of Borrelia burgdorferi and
  TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
  FILE REFERENCE: 2631.1002-001
  CURRENT APPLICATION NUMBER: US/09/596,746
  CURRENT FILING DATE: 2000-06-16
  PRIOR APPLICATION NUMBER: US 60/140,042
  PRIOR FILING DATE: 1999-06-18
  NUMBER OF SEQ ID NOS: 84
  SOFTWARE: FastSEQ for Windows Version 4.0
  SEQ ID NO 28
  LENGTH: 378
  TYPE: PRT
  ORGANISM: Artificial Sequence
  FEATURE:
  OTHER INFORMATION: OspC Chimera
US-09-596-746-28

Query Match    98.3%;   Score 1810.5;   DB 19;   Length 378;
Best Local Similarity  98.1%;   Pred. No. 1.7e-122;
Matches  371;   Conservative  2;   Mismatches  2;   Indels  3;   Gaps  1;

Qy    1  MACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKA  60
Db    1  MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA  60
Qy   61  IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN 120
Db   61  IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN 120
Qy  121  KLKEKHTDLGKEGYTDADAKEAILKNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK  180
Db  121  KLKEKHTDLGKEGYTDADAKEAILKANGTRTKGAEELGKLFESVEVLSKAAKEMLANSVK 180
Qy  181  ELTSPVVAES---PAMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEV 237
Db  181  ELTSPVVAESPKKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEV 240
Qy  238  ETLLTSIDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKA 297
Db  241  ETLLTSIDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKA 300
Qy  298  EIEKAKKCSEEFTAKLKGEHTDLGKEGYTDDNAKKAILKTNNDKTKGADELEKLFESVKN 357
Db  301  EIEKAKKCSEEFTAKLKGEHTDLGKEGYTDDNAKKAILKTNNDKTKGADELEKLFESVKN 360
Qy  358  LSKAAKEMLTNSVKELTS 375
Db  361  LSKAAKEMLTNSVKELTS 378

RESULT 8
US-09-596-746A-28
Sequence 28, Application US/0959674G A
GENERAL INFORMATION:
  APPLICANT: Dattwyler, Raymond J.
  APPLICANT: Seinost, Gerald
  APPLICANT: Dykhuizen, Daniel
  APPLICANT: Luft, Benjamin J.
  APPLICANT: Maria J.C. Gomes-Solecki
  TITLE OF INVENTION: Groups of Borrelia burgdorferi and
  TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
  FILE REFERENCE: 2631.1002-001
  CURRENT APPLICATION NUMBER: US/09/596,746A
  CURRENT FILING DATE: 2000-06-19
  PRIOR APPLICATION NUMBER: US 60/140,042
  PRIOR FILING DATE: 1999-06-18
  NUMBER OF SEQ ID NOS: 84
  SOFTWARE: FastSEQ for Windows Version 4.0
  SEQ ID NO 28
  LENGTH: 378
  TYPE: PRT
  ORGANISM: Artificial Sequence
  FEATURE:
  OTHER INFORMATION: OspC Chimera
US-09-596-746A-28

Query Match    98.3%;   Score 1810.5;   DB 19;   Length 378;
Best Local Similarity  98.1%;   Pred. No. 1.7e-122;
Matches  371;   Conservative  2;   Mismatches  2;   Indels  3;   Gaps  1;

Qy    1  MACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKA  60
Db    1  MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA  60
Qy   61  IGKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN 120
Db   61  IGKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN 120
Qy  121  KLKEKHTDLGKEGYTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 180
Db  121  KLKEKHTDLGKEGYTDADAKEAILKANGTRTKGAEELGKLFESVEVLSKAAKEMLANSVK 180
Qy  181  ELTSPVVAES---PAMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEV 237
Db  181  ELTSPVVAESPKKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEV 240
Qy  238  ETLLTSIDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKA 297
Db  241  ETLLTSIDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKA 300
Qy  298  EIEKAKKCSEEFTAKLKGEHTDLGKEGYTDDNAKKAILKTNNDKTKGADELEKLFESVKN 357
Db  301  EIEKAKKCSEEFTAKLKGEHTDLGKEGYTDDNAKKAILKTNNDKTKGADELEKLFESVKN 360
Qy  358  LSKAAKEMLTNSVKELTS 375
Db  361  LSKAAKEMLTNSVKELTS 378
```

RESULT 9
US-09-596-746A-26
Sequence 26, Application US/0959674 6A
GENERAL INFORMATION:
  APPLICANT: Dattwyler, Raymond J.
  APPLICANT: Seinost, Gerald
  APPLICANT: Dykhuizen, Danial
  APPLICANT: Luft, Benjamin J.
  APPLICANT: Maria J.C. Gomes-Solecki
  TITLE OF INVENTION: Groups of Borrelia burgdorferi and
  TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
  FILE REFERENCE: 2631.1002-001
  CURRENT APPLICATION NUMBER: US/09/596,746A
  CURRENT FILING DATE: 2000-06-19
  PRIOR APPLICATION NUMBER: US 60/140,042
  PRIOR FILING DATE: 1999-06-18
  NUMBER OF SEQ ID NOS: 84
  SOFTWARE: FastSEQ for Windows Version 4.0

  SEQ ID NO 26
  LENGTH: 374
  TYPE: PRT
  ORGANISM: Artificial Sequence
  FEATURE:
  OTHER INFORMATION: OspC Chimera
US-09-596-746A-26

Query Match   86.7%; Score 1597.5; DB 19; Length 374;
Best Local Similarity 87.7%; Pred. No. 4.3e-107;
Matches 328; Conservative 20; Mismatches 25; Indels 1; Gaps 1;

Qy   1 MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA 60

Qy   61 IGKKIHQNNGLDTEYNHNGSLLAGAYAISTLIIKQKLDGLKNEGLKEKIDAAKKCSEFFTN 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 IGKKIHQNNGLDTEYNHNGSLLAGAYAISTLIIKQKLDGLKNEGLKEKIDAAKKCSEFFTN 120

Qy   121 KLKEKHFDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 KLKEKHFDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 180

Qy   181 ELTSPVVAESPAMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVETL 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 ELTSPVVAESPAMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVETL 240

Qy   241 LTSIDELA-KAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAEI 299
          ||||||| |||||||||||||||::::||:||| ||:|:::::||:::|| |:||:|
Db   241 LASIDELATKAIGKKIQQNGGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKEKI 300

Qy   300 EKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLS 359
          ||||||||||||:||| ||| |||||||||||||||||||||||:||||||||:|
Db   301 ENAKKCSEDFTKLEGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLA 360

Qy   360 KAAKEMLTNSVKEL 373
          |||||| |||||||
Db   361 KAAKEMLANSVKEL 374

RESULT 10
US-09-596-746-26
Sequence 26, Application US/0959674 6
GENERAL INFORMATION:
  APPLICANT: Dattwyler, Raymond J.
  APPLICANT: Seinost, Gerald
  APPLICANT: Dykhuizen, Danial
  APPLICANT: Luft, Benjamin J.
  APPLICANT: Maria J.C. Gomes-Solecki
  TITLE OF INVENTION: Groups of Borrelia burgdorferi and
  TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans

---

  FILE REFERENCE: 2631.1002-001
  CURRENT APPLICATION NUMBER: US/09/596,746
  CURRENT FILING DATE: 2000-06-16
  PRIOR APPLICATION NUMBER: US 60/140,042
  PRIOR FILING DATE: 1999-06-18
  NUMBER OF SEQ ID NOS: 84
  SOFTWARE: FastSEQ for Windows Version 4.0

  SEQ ID NO 26
  LENGTH: 373
  TYPE: PRT
  ORGANISM: Artificial Sequence
  FEATURE:
  OTHER INFORMATION: OspC Chimera
US-09-596-746-26

Query Match   86.5%; Score 1592.5; DB 19; Length 373;
Best Local Similarity 87.7%; Pred. No. 9.9e-107;
Matches 327; Conservative 20; Mismatches 25; Indels 1; Gaps 1;

Qy   2 ACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAI 61
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAI 60

Qy   62 GKKIHQNNGLDTEYNHNGSLLAGAYAISTLIIKQKLDGLKNEGLKEKIDAAKKCSEFFTNK 121
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 GKKIHQNNGLDTEYNHNGSLLAGAYAISTLIIKQKLDGLKNEGLKEKIDAAKKCSEFFTNK 120

Qy   122 LKEKHFDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 181
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 LKEKHFDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 180

Qy   182 LTSPVVAESPAMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVETLL 241
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db   181 LTSPVVAESPAMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLL 240

Qy   242 TSIDELA-KAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAEIE 300
          ||||||| |||||||||||:::::||:||| ||:|:::::||:::||:|||:||:|
Db   241 ASIDELATKAIGKKIQQNGGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKEKIE 300

Qy   301 KAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLSK 360
          ||||||||:||| ||| |:||| ||||||| ||||| |:|||||:||||||:||:|:
Db   301 NAKKCSEDFTKKLEGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLAK 360

Qy   361 AAKEMLTNSVKEL 373
          |||||| ||||||
Db   361 AAKEMLANSVKEL 373

RESULT 11
US-09-596-746-62
Sequence 62, Application US/0959674 6
GENERAL INFORMATION:
  APPLICANT: Dattwyler, Raymond J.
  APPLICANT: Seinost, Gerald
  APPLICANT: Dykhuizen, Danial
  APPLICANT: Luft, Benjamin J.
  APPLICANT: Maria J.C. Gomes-Solecki
  TITLE OF INVENTION: Groups of Borrelia burgdorferi and
  TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
  FILE REFERENCE: 2631.1002-001
  CURRENT APPLICATION NUMBER: US/09/596,746
  CURRENT FILING DATE: 2000-06-16
  PRIOR APPLICATION NUMBER: US 60/140,042
  PRIOR FILING DATE: 1999-06-18
  NUMBER OF SEQ ID NOS: 84
  SOFTWARE: FastSEQ for Windows Version 4.0
  SEQ ID NO 62
  LENGTH: 397
  TYPE: PRT
  ORGANISM: OspC Chimera
US-09-596-746-62

```
Query Match          86.3%;  Score 1589.5;  DB 19;  Length 397;
Best Local Similarity 87.4%;  Pred. No. 1.8e-106;
Matches 326;  Conservative 21;  Mismatches 25;  Indels 1;  Gaps 1;

Qy    2 ACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAI   61
          :|||||||||||||||||||||||||||||||||||||||:||:|||||||||||||||||
Db   25 SCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAI   84

Qy   62 GKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAARKCSETFTNK  121
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   85 GKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAARKCSETFTNK  144

Qy  122 LKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE  181
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  145 LKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE  204

Qy  182 LTSPVVAESPAMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLL  241
          ||||||||||||||||||||||||||||||||||||||||||||:::||||||||||||
Db  205 LTSPVVAESPAMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLL  264

Qy  242 TSIDELA-KAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAEIE  300
          :||||| ::|||||| :::|| ::::|| ||:|:|||::::|| ::| :: ::|:|:||
Db  265 ASIDELATKAIGKKIQONGGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKEKIE  324

Qy  301 KAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLSK  360
          :|||||::||:||:|||:||:|||||:|||||||::: :||:||:||||||||::::|::
Db  325 NAKKCSEDFTKKLEGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLAK  384

Qy  361 AAKEMLTNSVKEL  373
          ||||||:|||||
Db  385 AAKEMLANSVKEL  397
```

RESULT 12
US-09-596-746A-62
; Sequence 62, Application US/0959746A
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seinost, Gerald
; APPLICANT: Dykhuizen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746A
; CURRENT FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 62
; LENGTH: 398
; TYPE: PRT
; ORGANISM: ospC Chimera
US-09-596-746A-62

```
Query Match          86.3%;  Score 1589.5;  DB 19;  Length 398;
Best Local Similarity 87.4%;  Pred. No. 1.8e-106;
Matches 326;  Conservative 21;  Mismatches 25;  Indels 1;  Gaps 1;

Qy    2 ACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAI   61
          :|||||||||||||||||||||||||||||||||||||||:||:|||||||||||||||||
Db   26 SCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAI   85

Qy   62 GKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAARKCSETFTNK  121
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   86 GKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAARKCSETFTNK  145

Qy  122 LKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE  181
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  146 LKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE  205

Qy  182 LTSPVVAESPAMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLL  241
          ||||||||||||||||||||||||||||||||||||||||||||:::||||||||||||
Db  206 LTSPVVAESPAMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLL  265

Qy  242 TSIDELA-KAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAEIE  300
          :||||| ::|||||| :::|| ::::|| ||:|:|||::::|| ::| :: ::|:|:||
Db  266 ASIDELATKAIGKKIQONGGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKEKIE  325

Qy  301 KAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLSK  360
          :|||||::||:||:|||:||:|||||:|||||||::: :||:||:||||||||::::|::
Db  326 NAKKCSEDFTKKLEGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLAK  385

Qy  361 AAKEMLTNSVKEL  373
          ||||||:|||||
Db  386 AAKEMLANSVKEL  398
```

RESULT 13
US-09-596-746A-82
; Sequence 82, Application US/0959746
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seinost, Gerald
; APPLICANT: Dykhuizen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746
; CURRENT FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 82
; LENGTH: 400
; TYPE: PRT
; ORGANISM: ospC Chimera
US-09-596-746A-82

```
Query Match          85.5%;  Score 1575;  DB 19;  Length 400;
Best Local Similarity 86.4%;  Pred. No. 2e-105;
Matches 325;  Conservative 22;  Mismatches 25;  Indels 4;  Gaps 2;

Qy    2 ACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAI   61
          :|||||||||||||||||||||||||||||||||||||||:||:|||||||||||||||||
Db   25 SCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAI   84

Qy   62 GKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAARKCSETFTNK  121
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   85 GKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAARKCSETFTNK  144

Qy  122 LKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE  181
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  145 LKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE  204

Qy  182 LTSPVVAES---PAMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVE  238
          |||||||||   :||||||||||||||||||||||||||||||||||||||||||||||:
Db  205 LTSPVVAESPKKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIE  264

Qy  239 TLLTSIDELA-KAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKA  297
          |||:||||| ::|||||| :::|| ::::|| ||:|:|||::::|| ::| :: ::|:
Db  265 TLLASIDELATKAIGKKIQONGGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKE  324

Qy  298 EIEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKN  357
          :||:|||||::||:||:|||:||:|||||:|||||||::: :||:||:||||||||::::
Db  325 KIENAKKCSEDFTKKLEGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVEN  384

Qy  358 LSKAAKEMLTNSVKEL  373
          |||||||||:|||||
Db  325 LSKAAKEMLANSVKEL  373
```

```
         |:||||||| ||||||
Db  385 LAKAAKEMLANSVKEL 400
```

RESULT 14
US-09-596-746A-82
; Sequence 82, Application US/0959746A
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Selnost, Gerald
; APPLICANT: Dykhuizen, Danial
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
;   Borrelia afzelii That Cause Lyme Disease in Humans
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746A
; CURRENT FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 82
; LENGTH: 401
; TYPE: PRT
; ORGANISM: ospC Chimera
US-09-596-746A-82

Query Match   85.5%; Score 1575; DB 19; Length 401;
Best Local Similarity 86.4%; Pred. No. 2e-105;
Matches 325; Conservative 22; Mismatches 25; Indels 4; Gaps 2;

Qy   2 ACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAI 61
         |||||||||||||||||||||||||||||||||||||||||||||||||||||:||
Db   26 SCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAI 85

Qy   62 GKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNK 121
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   86 GKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNK 145

Qy 122 LKEKHTDLGKEGYTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 181
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 146 LKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 205

Qy 182 LITSPVVAES---PAMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVE 238
         ||||||| : | |:|||||||||||||||||||||||||||||||||||||||||
Db 206 LITSPVVAESPKKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIE 265

Qy 239 TLLTSIDELA-KAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKA 297
         ||| ||||| |||||:|| : | | ||:|:||| || ||:|||||||||: : ||
Db 266 TLLASIDELATKAIGKKIQONGGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKE 325

Qy 298 EIEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKN 357
         :|| |||| : |:|| ||||||||||||||||||||||||| ||| ||||||||:|:|
Db 326 KIENAKKCSEDFTKKLEGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVEN 385

Qy 358 LSKAAKEMLTNSVKEL 373
         ||:|||||| ||||||
Db 386 LAKAAKEMLANSVKEL 401
```

RESULT 15
US-09-596-746-30
; Sequence 30, Application US/09596746
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Selnost, Gerald
; APPLICANT: Dykhuizen, Danial
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
;   Borrelia afzelii That Cause Lyme Disease in Humans

; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746
; CURRENT FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 377
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: OspC Chimera
US-09-596-746-30

Query Match      85.0%;  Score 1566;  DB 19;  Length 377;
Best Local Similarity  85.9%;  Pred. No. 8.3e-105;
Matches 324;  Conservative 22;  Mismatches 27;  Indels 4;  Gaps 2;

Qy    1 MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA 60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db    1 MACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKA 60

Qy   61 IGKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN 120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN 120

Qy  121 KLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 KLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 180

Qy  181 ELITSPVVAES---PAMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEV 237
         |||||||| : |  |:|||||||||||||||||||||||||||||||||||||||||
Db  181 ELITSPVVAESPKKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEI 240

Qy  238 ETLLTSIDELA-KAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELK 296
         || ||||||| |||||:||: | || ||:|:|||| || ||:|||||||||: : ||
Db  241 ETLLASIDELATKAIGKKIQONGGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLK 300

Qy  297 AEIEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVK 356
         :|| |||||: |:||  ||||||||||||||||||||||||| ||| |||||||||:|:
Db  301 EKIENAKKCSEDFTKKLEGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVE 360

Qy  357 NLSKAAKEMLTNSVKEL 373
         ||:|||||| ||||||
Db  361 NLAKAAKEMLANSVKEL 377
```

Search completed: March 18, 2002, 10:08:42
Job time: 970 sec

OM protein - protein search, using sw model

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

Run on: March 18, 2002, 09:52:02 ; Search time 81.02 Seconds
               (without alignments)
               295.677 Million cell updates/sec

Title: US-09-596-746A-24
Perfect score: 1842
Sequence: 1 MACNNSGKDGNTSANSADES..........KNLSKAAKEMLTNSVKELTS 375

Scoring table: BLOSUM62
Searched: Gapop 10.0 , Gapext 0.5
               303921 seqs, 63882009 residues

Database :   Pending_Patents_AA_New:*
               1: /cgn2_6/ptcdata/2/paa/PCT_NEW_COMB.pep:*
               2: /cgn2_6/ptcdata/2/paa/US06_NEW_COMB.pep:*
               3: /cgn2_6/ptcdata/2/paa/US07_NEW_COMB.pep:*
               4: /cgn2_6/ptcdata/2/paa/US08_NEW_COMB.pep:*
               5: /cgn2_6/ptcdata/2/paa/US09_NEW_COMB.pep:*
               6: /cgn2_6/ptcdata/2/paa/US10_NEW_COMB.pep1:*
               7: /cgn2_6/ptcdata/2/paa/US10_NEW_COMB.pep:*
               8: /cgn2_6/ptcdata/2/paa/US60_NEW_COMB.pep:*

               Listing first 45 summaries

Post-processing: Minimum Match 0%
                 Maximum Match 100%

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Total number of hits satisfying chosen parameters: 303921

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	%	Query Match Length	DB	ID	Description
1	674.5	36.6	211	6	US-09-974-992-5	Sequence 5, Appli
2	614	33.3	212	6	US-09-974-992-7	Sequence 7, Appli
3	587.5	31.9	207	6	US-10-072-851-5835	Sequence 3, Appli
4	175	9.5	1014	6	US-09-708-427-19883	Sequence 3, Appli
5	175	9.5	1018	6	US-09-708-427-19882	Sequence 3, Appl1
6	175	9.5	1269	6	US-09-708-427-19881	Sequence 1, Appl1
7	169.5	9.2	1313	6	US-09-708-427-15044	Sequence 15044, A
8	169	9.2	1881	6	US-10-032-585-7646	Sequence 7646, Ap
9	169	9.2	1881	7	US-10-072-851-15590	Sequence 15590, A
10	168.5	9.1	1304	6	US-09-708-427-15045	Sequence 15045, A
11	158	8.6	589	6	US-09-748-875-14	Sequence 14, Appl
12	158	8.6	690	5	US-09-748-875-61	Sequence 61, Appl
13	158	8.6	691	5	US-09-748-875-1	Sequence 1, Appl
14	158	8.6	701	5	US-09-748-875-62	Sequence 62, Appl
15	158	8.6	707	5	US-09-748-875-2	Sequence 2, Appli
16	158	8.6	837	7	US-09-815-242-5883	Sequence 5883, Ap
17	158	8.6	837	7	US-10-072-851-5883	Sequence 5883, Ap
18	158	8.6	875	7	US-09-815-242-13080	Sequence 13080, A
19	158	8.6	875	7	US-10-072-851-13080	Sequence 13080, A
20	158	8.6	929	7	US-09-748-875-60	Sequence 60, Appl
21	157.5	8.6	1725	7	US-10-037-182-12	Sequence 12, Appl.
22	157.5	8.6	1786	5	US-09-938-275-7	Sequence 7, Appl1
23	157.5	8.6	1786	7	US-10-037-182-10	Sequence 10, Appl
24	154.5	8.4	1144	7	US-09-708-427-15046	Sequence 15046, A
25	154.5	8.4	2434	6	US-09-815-242-5835	Sequence 5835, Ap
26	154.5	8.4	2434	7	US-10-072-851-5835	Sequence 5835, Ap
27	154.5	8.4	211	6	US-09-815-242-12996	Sequence 12996, A
28	154.5	8.4	212	6	US-10-072-851-12996	Sequence 12996, A
29	152.5	8.3	6281	7	US-09-708-427-21161	Sequence 21161, A
30	152.5	8.3	1703	6	US-09-708-427-21160	Sequence 21160, A
31	152.5	8.3	1710	6	US-09-708-427-21159	Sequence 21159, A
32	152.5	8.3	1744	6	US-09-708-427-21159	Sequence 21159, A
33	152.5	8.3	2437	6	US-10-072-851-5834	Sequence 5834, Ap
34	152.5	8.2	2437	7	US-09-815-242-5834	Sequence 5834, Ap
35	150.5	8.2	670	5	US-09-748-875-63	Sequence 63, Appl
36	150	8.2	1690	6	US-09-748-875-10224	Sequence 10224, A
37	150	8.1	805	6	US-09-614-150-10311	Sequence 10311, A
38	150	8.1	841	6	US-09-614-150-10224	Sequence 10224, A
39	150	7.9	711	5	US-10-037-182-8	Sequence 8, Appl1
40	146	7.9	1765	7	US-10-037-182-6	Sequence 6, Appl1
41	146	7.9	1786	5	US-09-708-427-3730	Sequence 3730, Ap
42	146	7.9	1786	6	US-09-938-275-6	Sequence 6, Appl1
43	146	7.9	1786	7	US-09-561-709B-9	Sequence 9, Appl1
44	146	7.9	1786	7	US-10-037-182-6	Sequence 6, Appl1
45	146	7.9	2086	6	US-09-815-242-5639	Sequence 5639, Ap

ALIGNMENTS

RESULT 1
US-09-974-992-5

Sequence 5, Application US/09974992

GENERAL INFORMATION:
APPLICANT: Mathiesen, Marianne J.
APPLICANT: Theisen, Michael
APPLICANT: Holm, Arne
APPLICANT: Ostergaard, Soren
TITLE OF INVENTION: Novel OspC-derived peptide fragments
FILE REFERENCE: 459-666P
CURRENT APPLICATION NUMBER: US/09/974,992
CURRENT FILING DATE: 2001-10-10
PRIOR APPLICATION NUMBER: US/09/180,089
PRIOR FILING DATE: 1999-05-13
PRIOR APPLICATION NUMBER: PCT/DK97/00203
PRIOR FILING DATE: 1997-05-02
NUMBER OF SEQ ID NOS: 40
SOFTWARE: Patentin Ver. 2.1

SEQ ID NO 5
LENGTH: 211
TYPE: PRT
ORGANISM: Borrelia burgdorferi
US-09-974-992-5

Query Match   36.6%; Score 674.5; DB 6; Length 211;
Best Local Similarity   76.7%; Pred. No. 2.4e-39;
Matches 148; Conservative   11; Mismatches   31; Indels   3; Gaps   3;

```
Qy   1 MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA 60
         |||||  |||||||||||||||||| | ||||||||||||||||||||||||:  ||||
Db  17 ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEL-AKA 75

Qy  61 IGKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLK-NEGLKEKIDAAKKCSETFT 119
         |||||  | | ||||||| |||||||||||||||||| |  |||||||||||| :  ||
Db  76 IGKKIKNDGSLGDEANHNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEEFS 135

Qy 120 NKLKEKHTDLGKEGVTDADAKEAILKTNGT-KTKGAEELGKLFESVEVLSKAAKEMLANS 178
         |||  | ||| ||||| :|||:|||| |  || || |||  ||||||||||||||||
Db 136 TKLKDNHAQLGIQGVTDENKKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANS 195

Qy 179 VKELTSPVVAESP 191
         |||||||||||||
Db 196 VKELTSPVVVESP 208
```

RESULT 2
US-09-974-992-7

Sequence 7, Application US/09974992
GENERAL INFORMATION:
- APPLICANT: Mathiesen, Marianne J.
- APPLICANT: Theisen, Michael
- APPLICANT: Holm, Arne
- APPLICANT: Ostergaard, Soren
- TITLE OF INVENTION: Novel OspC-derived peptide fragments
- FILE REFERENCE: 459-666P
- CURRENT APPLICATION NUMBER: US/09/974,992
- CURRENT FILING DATE: 2001-10-10
- PRIOR APPLICATION NUMBER: 09/180,089
- PRIOR FILING DATE: 1999-05-13
- PRIOR APPLICATION NUMBER: PCT/DK97/00203
- PRIOR FILING DATE: 1997-05-02
- NUMBER OF SEQ ID NOS: 40
- SOFTWARE: Patentin Ver. 2.1
- SEQ ID NO 7
- LENGTH: 212
- TYPE: PRT
- ORGANISM: Borrelia afzelii

RESULT 3
US-09-974-992-3

Query Match
Best Local Similarity 33.3%; Score 614; DB 6; Length 212;
Matches 135; Conservative 14; Mismatches 42; Indels 2; Gaps 2;

```
Qy   1 MACNNSGKDGNT-SANSADESVKGPNLTEISKKIITDSNAVLLAVKEVEALLSSIDEIAAK 60
         ::|||||  :|||||||| : |||  |||||||||| :: ||||||||||||||:: ||
Db  17 ISCNNSG--GDSASTNPDESAKGPNLTEISKKIITDSNAFLVAVKEVETLVLSIDELAKK 76

Qy  60 AIGKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKN-EGLKEKHIDAAKKCSETF 118
        |||:| :||||||||||||||||||||||||:| | | ||  ||||||| || |||||||
Db  77 AIGKKIDNNGLLAALNNQNGSLLAGAYAISTLITEKLSKIKNLEELKTEIAKAKKCSEEF  136

Qy 119 TNKLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANS 178
        |||:| | : | |||:||||||| | | |||||:|||:||||||||||:|||| ||  |
Db 137 TNKLKSGHADLGKQDATDDHAKAAILKTHATIDKGAKEFKDLFESVEGLLKAAQVALTNS 196

Qy 179 VKELTSPVVAESP 191
        |||||||||||||
Db 197 VKELTSPVVAESP 209
```

Sequence 3, Application US/09974992
GENERAL INFORMATION:
- APPLICANT: Mathiesen, Marianne J.
- APPLICANT: Theisen, Michael
- APPLICANT: Holm, Arne
- APPLICANT: Ostergaard, Soren
- TITLE OF INVENTION: Novel OspC-derived peptide fragments
- FILE REFERENCE: 459-666P
- CURRENT APPLICATION NUMBER: US/09/974,992
- CURRENT FILING DATE: 2001-10-10
- PRIOR APPLICATION NUMBER: 09/180,089
- PRIOR FILING DATE: 1999-05-13
- PRIOR APPLICATION NUMBER: PCT/DK97/00203
- PRIOR FILING DATE: 1997-05-02
- NUMBER OF SEQ ID NOS: 40
- SOFTWARE: Patentin Ver. 2.1
- SEQ ID NO 3
- LENGTH: 207
- TYPE: PRT
- ORGANISM: Borrelia garinii

Query Match
Best Local Similarity 31.9%; Score 587.5; DB 6; Length 207;
Matches 127; Conservative 24; Mismatches 37; Indels 3; Gaps 2;

RESULT 4
US-09-708-427-19883
Sequence 19883, Application US/09708427
GENERAL INFORMATION:
- APPLICANT: N. ALEXANDROV et al.
- TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
- TITLE OF INVENTION: THEREBY
- FILE REFERENCE: 2750-1243P
- CURRENT APPLICATION NUMBER: US/09/708,427
- CURRENT FILING DATE: 2000-11-09
- NUMBER OF SEQ ID NOS: 85364
- SOFTWARE: Patentin version 3.1
- SEQ ID NO 19883
- LENGTH: 1014
- TYPE: PRT
- ORGANISM: Arabidopsis thaliana
- FEATURE:
- NAME/KEY: misc_feature
- LOCATION: 1..1014
- OTHER INFORMATION: Xaa is any amino acid
- NAME/KEY: misc_feature
- LOCATION: 1..1014
- OTHER INFORMATION: Ceres Seq. ID 1836847

Query Match
Best Local Similarity 9.5%; Score 175; DB 6; Length 1014;
Matches 118; Conservative 70; Mismatches 161; Indels 22; Gaps 22;

```
Qy   8 KDGNISANSADE----------------SVKGPNL-------TEISKKIITDS-NAVLLAVKE 45
        :|| :  | |                   |  ||:        ||:|||||:  ||::||||:
Db 401 KHGETEADSKGYLGQVAELQSTLEAFQVKSSSLEAALNIATENEKELTENLNAVTSEKKK   460

Qy  46 VEALLSSIDEIAAK------AIGKKIHQNNGLDTEYNHNGSLLAGAYAISTLI--IKQ 94
        :|| | : |  :|        | :  | :   |   :|  | |  |  | :  | ||
Db 461 LEA----TVDEYSVKISESENLLESIRNELNVTQG-KLESIENDLKAAGLQESEVMEKLKS  516

Qy  95 KLDGLKNEGLKEKIDAA----------------KKCSETFTNK-------LKEKH 126
        :| |:||| |||:|:                  ||::::| |        |:|||
Db 517 AEESLEQKG--REIDEATTKRMELEALHQSLSIDSEHRLQKAMEEFTSRDSEASSLTEKL  574

Qy 127 TDL-GKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKELTSP 185
         :| | :    : :  ::|  ::   |       |  :  ::     :  :|: :|  | |
Db 575 RDLEGKIKSYEEQLAEASGKRSSSLKEKLEQTLGRLAAAESVNEKLKQEEDQAQEKSLQSS 634

Qy 186 VVAESPAMVNNSGK---------DGNTSANSADESVKGPNLTE-----ISKKITESNAVVLAVK 235
         : | :   ::|:          |||:||: : :|:|:  |:      :||| | ||||| :
Db 635 SESELLAETNNQLKIKIQELEGLIGSGSVEKETALKRLEEAIERFNQKETESSSDLVEKLK  694

Qy 236 EVETLLTSIDELA-KAIG-----------------KKIKN---------------- 257
        :| :| ::|| |  ||:                  ||:||
Db 695 THENQIEEYKKLAHEASGVADTRKVELEDALSKIKNLESTIEELGAKCQQGLEKESGDLAE  754

Qy 258 -DVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKA-EIEKAKKCSEEFTAKL-- 313
         :: |: :| ||:|:|:|| |||:| |:|:|:||:|| |:  |:|:|||||||||| |
```

```
Db   755  VNLKLNLELANHGS------------EANELQTKLSALEAEKEQTANELEARASKTTIEDLTKQLTS  807
          : ||                      :| :| | ||| :| ||:| |: ::
Qy   314  KGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVK-NLSKAAKEMLTNSVK-  371

Db   808  EGEKLQSOIEKLRAVAAEKSVLESH------FEELEKTLSEVKAQLKENVENAATASVKV  861
          |||| :| | :  |  |:|:              ||:| ||| || |:|

Qy   372  -ELTS  375
           ||||
Db   862  AELTS  866
```

RESULT   5

US-09-708-427-19882
; Sequence 19882, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: THEREBY
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 19882
; LENGTH: 1018
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..1018
; OTHER INFORMATION: Xaa is any amino acid
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..1018
; OTHER INFORMATION: Ceres Seq. ID 1836846

US-09-708-427-19882

Query Match   9.5%; Score 175; DB 6; Length 1018;
Best Local Similarity   24.3%; Pred. No. 0.0004;
Matches 118; Conservative 70; Mismatches 161; Indels 136; Gaps 22;

```
Qy     8  KDGNTSANSADE------------SVKGPNL------------TEISKKITDS-NAVLLAVKE   45
          | |:|                  :|                :||| :|| |||  :
Db   405  KHGETEADSKGYLGQVAELQSTLEAFQVKSSSLEAALNIATENEKELTENLNAVTSEKKK  464

Qy    46  VEALLSSIDEIAAK------------AIGKKIHQNNGLDTEYNHNGSLLAGAYAISTL--IKQ   94
          :||| :|| ||               ::| |: | :||  :| | |
Db   465  LEA---TVDEYSVKISESENLLESIRNELNVTQG-KLESIENDLKAAGLQESEVMEKLKS  520

Qy    95  KLDGLKNEGLKEKIDAA------------KKCSEFFTNK-----LKEKH  126
          ::|:| |                      ||| :|| ||      ||||
Db   521  AEESLEOKG--REIDEATTKRMELEALHQSLSIDSEHRLQKAMEEFTSRDSEASSLTEKL  578

Qy   127  TDL-GKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKELTSP  185
          |:| ||:|| ||  | | |  | ||:|| |: | | |  ||||:| ||||  :||
Db   579  RDLEGKIKSYEEQLAEASGKSSSLKEKLEQTLGRLAAAESVNEKLKQEFDQAOERSLQSS  638

Qy   186  VVAESPAMVNNSGK------------DGNTSANSADESVKGPNLTE---  235
          :   |   ::  |:            ::  :||  :|
Db   639  SESELLAETNNQLKIKIQELEGLIGSGSVEKETALKRLEEAIERENQKETESSDLVEKLK  698

Qy   236  EVETLLTSIDELA-KAIG------------KKIKN  257
          :|  |:|:|  |  ||||            ||:|
Db   699  THENQIEEYKKLAHEASGVADTRKVELEDALSKLKNLESTIEELGAKCQGLEKESGDLAE  758

Qy   258  -DVSLDNEADHNGSLISGAYLISNLITTKKISAIKDSGELKA-EIEKAKKCSEEFTAKL--  313
           :: | :  ||  |:|| | :|||  || |  ||:  |||  :|:|| |
Db   759  VNLKLNLELANHGS------------EANELQTKLSALEAEKEQTANELEARASKTTIEDLTKQLTS  811

Qy   314  KGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVK-NLSKAAKEMLTNSVK-  371
          :| : ||                  |:| |:| |  ||:| |:  |:|  || | |:
Db   812  EGEKLQSOIEKLRAVAAEKSVLESH------FEELEKTLSEVKAQLKENVENAATASVKV  865
```

```
Qy   372  -ELTS  375
           ||||
Db   866  AELTS  870
```

RESULT   6

US-09-708-427-19881
; Sequence 19881, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
; TITLE OF INVENTION: THEREBY
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 19881
; LENGTH: 1269
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..1269
; OTHER INFORMATION: Xaa is any amino acid
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..1269
; OTHER INFORMATION: Ceres Seq. ID 1836845

US-09-708-427-19881

Query Match   9.5%; Score 175; DB 6; Length 1269;
Best Local Similarity   24.3%; Pred. No. 0.00053;
Matches 118; Conservative 70; Mismatches 161; Indels 136; Gaps 22;

```
Qy     8  KDGNTSANSADE------------SVKGPNL------------TEISKKITDS-NAVLLAVKE   45
          | |:|                  :|                :||| :|| |||  :
Db   656  KHGETEADSKGYLGQVAELQSTLEAFQVKSSSLEAALNIATENEKELTENLNAVTSEKKK  715

Qy    46  VEALLSSIDEIAAK------------AIGKKIHQNNGLDTEYNHNGSLLAGAYAISTL--IKQ   94
          :||| :|| ||               ::| |: | :||  :| | |
Db   716  LEA---TVDEYSVKISESENLLESIRNELNVTQG-KLESIENDLKAAGLQESEVMEKLKS  771

Qy    95  KLDGLKNEGLKEKIDAA------------KKCSEFFTNK-----LKEKH  126
          ::|:| |                      ||| :|| ||      ||||
Db   772  AEESLEOKG--REIDEATTKRMELEALHQSLSIDSEHRLQKAMEEFTSRDSEASSLTEKL  829

Qy   127  TDL-GKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKELTSP  185
          |:| ||:|| ||  | | |  | ||:|| |: | | |  ||||:| ||||  :||
Db   830  RDLEGKIKSYEEQLAEASGKSSSLKEKLEQTLGRLAAAESVNEKLKQEFDQAOERSLQSS  889

Qy   186  VVAESPAMVNNSGK------------DGNTSANSADESVKGPNLTE---  235
          :   |   ::  |:            ::  :||  :|
Db   890  SESELLAETNNQLKIKIQELEGLIGSGSVEKETALKRLEEAIERENQKETESSDLVEKLK  949

Qy   236  EVETLLTSIDELA-KAIG------------KKIKN  257
          :|  |:|:|  |  ||||            ||:|
Db   950  THENQIEEYKKLAHEASGVADTRKVELEDALSKLKNLESTIEELGAKCQGLEKESGDLAE  1009

Qy   258  -DVSLDNEADHNGSLISGAYLISNLITTKKISAIKDSGELKA-EIEKAKKCSEEFTAKL--  313
           :: | :  ||  |:|| | :|||  || |  ||:  |||  :|:|| |
Db  1010  VNLKLNLELANHGS------------EANELQTKLSALEAEKEQTANELEARASKTTIEDLTKQLTS  1062

Qy   314  KGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVK-NLSKAAKEMLTNSVK-  371
          :| : ||                  |:| |:| |  ||:| |:  |:|  || | |:
Db  1063  EGEKLQSOIEKLRAVAAEKSVLESH------FEELEKTLSEVKAQLKENVENAATASVKV  1116

Qy   372  -ELTS  375
           ||||
Db  1117  AELTS  1121
```

RESULT   7

US-09-708-427-15044
Sequence 15044, Application US/09708427
GENERAL INFORMATION:
APPLICANT: N. ALEXANDROV et al.
TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
TITLE OF INVENTION: THEREBY
FILE REFERENCE: 2750-1243P
CURRENT APPLICATION NUMBER: US/09/708,427
CURRENT FILING DATE: 2000-11-09
NUMBER OF SEQ ID NOS: 85364
SOFTWARE: Patentin version 3.1
SEQ ID NO 15044
LENGTH: 1313
TYPE: PRT
ORGANISM: Arabidopsis thaliana
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1..1313
OTHER INFORMATION:
NAME/KEY: misc_feature
LOCATION: 1..1313
OTHER INFORMATION: Xaa is any amino acid
US-09-708-427-15044

Query Match
Best Local Similarity   22.2%;
Matches 118;   Conservative 60;   Mismatches 140;   Indels 213;   Gaps 23;
Score 169.5;   DB 6;   Pred. No. 0.0013;   Length 1313;

```
Db    8  KDGNTSANSADESVKGPNLTEISKKITDSNAV-
         :::|||
Qy    6  KTGLMETPRSKPSPPPRLSKLSASKSDSNSASPKVHSRLVKGTELQTQLNQIQEDLKKA  65
                                                :::|||  ||
Db    6  ------------------------------LLAVKEVEALLSSIDEIAAKA  60

Qy   61  IGKKIHQNNGLDTEYNHNGSLLAGAYAISTLLIKQK---LDGLK-------NEGLKEKI 108
Db   66  -----DEQ-------IELLKKDKAKAIDDLIKESEKLVEEANEKLKEAL 101

Qy  109  DAAKKCSETF----TNK----LKEKHT-DLG-------- 130
Db  102  AAQKRAEESFEVEKFRAVELEQAGLEAVOKKDVTSKNELESIRSQHALDISALLSTTEEL 161

Qy  131  ---KEGVTDADAKEAILKTNGTKTKGAE------ELGKL--------FES 163
Db  162  QRVKHELSMTADAKNKALSHAEEATKIAEHAEKAEILASELGRIKALLGSKEEKPAIEG 221

Qy  164  VEVLSKAAKE--------MLANSVKE------LITSPVVAESPAMVNNSG-- 198
Db  222  NEIVSLKLKSEIELLRGELEKVSILESSLKEQEGLVEQLKVDLEAAKMAES---CTNSSVE 278

Qy  199  -------KDGNTSANSADESVKGPNLTEISKKITESNAVV--------- 231
Db  279  EWRNKVHELEKEVEESNRSKSSASESME-----SVMKQLAELNHVLHETKSDNAAQKEKI 333

Qy  232  -LAVKEVETLLTSIDE----LAKAIGKKIKNDV--------SLDNEADHN 268
Db  334  ELLEKTIEAQRIDLEEYGROVCIAKEEASKLENLVESIKSELEISQEEKTRALDNEKAAT 393

Qy  269  GSLISGAYLISNLIITKKISAIKDSGELKAEIEAKKCSEEFTAKLKGEHTDLGKEGVTDD 328
Db  394  SN------IQNLLDQRTELSIELERCKVEEEASKKKDMESLTLAL------QEASTES 438

Qy  329  NAKKAILKTNNDKTKGADELEKLFESVKNLSKAAKE--------MLTNSVKELTS 375
Db  439  SEAKATLLVCQEELKNC----ESQVDSLKLASKETNEKYEMLEDARNEIDS 486
```

RESULT 8
US-10-032-585-7646
Sequence 7646, Application US/10032585
GENERAL INFORMATION:
APPLICANT: Terry, Roemer D.
APPLICANT: Bo, Jiang
APPLICANT: Charles, Boone

APPLICANT: Howard, Bussey
TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
FILE REFERENCE: 10182-005-999
CURRENT APPLICATION NUMBER: US/10/032,585
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 8000
SOFTWARE: Patentin version 3.1
SEQ ID NO 7646
LENGTH: 1881
TYPE: PRT
ORGANISM: Candida albicans
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (1881)..(1881)
OTHER INFORMATION: X=any amino acid
US-10-032-585-7646

Query Match
Best Local Similarity   20.7%;
Matches 98;   Conservative 79;   Mismatches 172;   Indels 124;   Gaps 16;
Score 169;   DB 7;   Pred. No. 0.0023;   Length 1881;

```
Qy    1  MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA  60
Db 1411  LAAQLETKTSNLDSTTMELEKTELELKKVRNELTEATSELTKLQDDNNOSLTEEIEKTKAA 1470

Qy   61  IGKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKLKNE------GLKEKIDAAKKC 114
Db 1471  LTK---SSKDLEVCGNQKSELQDSLKS----VKSELKNFENKYNQETTSLKDEIEEKQKE 1523

Qy  115  SETFTNKLLKEKHTDLGKEGVTDDADAKEAILKT-----------NG 148
Db 1524  IVTLQTELKDRISEVEKERAMLSENSETVIKEYSDKIKSLESKINSIKENHSKEITTHNE 1583

Qy  149  TKTKGAEELGKLFESVEVLSKAAKEMLANSVKELTSPVVAESPAMVNNSGKDGNTSANSA 208
Db 1584  QKTSLKQDIAKLSQDHE-SAQTQLEDKENQLKELKASL---------EKIHTESATSI 1631

Qy  209  DESVKGPNLTEISKKI------------TESNAVVLAVKEVETLLTSIDELAKAIGK-----K 254
Db 1632  EE--KNNQIKELSETIKSLKTELKQTSGDALKQSOKEYKTLKTKNSDTESKLEKQLEELEK 1689

Qy  255  IKNDVSLDNEADHNGSLISGAYLISNLIITKKISAIKDSGELKA------EIEKAK--- 303
Db 1690  VKSDLQTADE-KLKGITERELALKSELETVKNGLSTTSELAALTKTVKSLEKEREELQF 1748

Qy  304  ----RCSEEFTAKLKGEHTDLGK--EGVTD------DNAKKAILKTNNDKTKGAD 346
Db 1749  LSGNKSKELEDYIQK---HSDISEKLKALIDELKEKTKQFDDSKKKLTELENDLTSTKK 1804

Qy  347  ELE-----------------KLFESVKNLSKAAKEMLTNSVKELTS 375
Db 1805  ELETEKTIQTSKFKNLEERKDKEIVKLNKELELLKNDNSGAKKELSEKVSKLES 1857
```

RESULT 9
US-10-072-851-15590
Sequence 15590, Application US/10072851
GENERAL INFORMATION:
APPLICANT: Carr, Grant J.
APPLICANT: Xu, H. Howard
APPLICANT: Foulkes, J. Gordon
APPLICANT: Zamudio, Carlos
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Roemer, Terry
APPLICANT: Jiang, Bo
APPLICANT: Boone, Charles
APPLICANT: Bussey, Howard

; TITLE OF INVENTION: Methods for Identifying the Target of a Compound which Inhibits C
; TITLE OF INVENTION: Proliferation
; FILE REFERENCE: ELITRA.028A
; CURRENT APPLICATION NUMBER: US/10/072,851
; CURRENT FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 15811
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15590
; LENGTH: 1881
; TYPE: PRT
; ORGANISM: Candida albicans
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (1881)..(1881)
; OTHER INFORMATION: X=any amino acid
US-10-072-851-15590

Query Match
Best Local Similarity 20.7%;
Matches 98; Conservative 79; Pred. No. 0.0023; Mismatches 172; Indels 124; Gaps 16;
Score 169; DB 7; Length 1881;

```
Qy    1 MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSSIDEIAAKA   60
Db 1411 LAAQLETKTSNLDSTTMELEKTELELKKVRNELTEATSELTKLQDNNQSLTEELTEKTKAA 1470

Qy   61 IGKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNE------GLKEKIDAAKKC  114
Db 1471 LTK--SSKDLEVCGNQKSELQDSLKS-----VKSELKNFENKINQETTSLKDEIEEKQKE 1523

Qy  115 SETFTNKLKEKHTDLGKEGVTDADAKEAILKT----------------NG           148
Db 1524 IVTLQTELKDRISEVEKEKRAMLSENSETVIKEYSDKIKSLESKINSIKENHSKEITTHNE 1583

Qy  149 TKTKGAEELGKLFESVEVLSKAAKEMLANSVKELTSPVVAESPAMVNNSGKDGNTSANSA  208
Db 1584 QKTSLKQDIAKLSQDHE-SAQTQLEDKENQLKELKASL------------EKHNTESATSI 1631

Qy  209 DESVKGPNLTEISKKI-----------TESNAVVLAVKEVETLLTSIDELAKAIGK----K 254
Db 1632 EE--KNNQIKELSETIKSLKTELKTSGDALKQSQKEYKTLKTKNSDTESKLEKQLEELEK  1689

Qy  255 IKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKA-----EIEKAK----- 303
Db 1690 VKSDLQTADE-KLKGITEREIALKSELETVKNSGLSTTSELAALTKTVKSLEKEKEELQF 1748

Qy  304 ------KCSEEFTAKLKGEHTDLGK--EGVTD--------DNAKKAILKTNNDKTKGAD  346
Db 1749 LSGNKSKELEDYIQK-----HSDISEKLKALTDELKEKTKQFDDSKKLTELENDLTSTKK 1804

Qy  347 ELE--------------------------KLFESVKNLSKAAKEMLTNSVKELTS      375
Db 1805 ELETEKTQTSKFKNLEERKDKEIVKLNKELELLKNDNGAAKKELSEKVSKLES        1857
```

RESULT 10
US-09-708-427-15045
Sequence 15045, Application US/09708427
GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 15045
; LENGTH: 1304
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..1304
; NAME/KEY: misc_feature
; LOCATION: 1..1304
; OTHER INFORMATION: Ceres Seq. ID 1828628
US-09-708-427-15045

Query Match
Best Local Similarity 22.3%;
Matches 115; Conservative 59; Pred. No. 0.0015; Mismatches 128; Indels 213; Gaps 23;
Score 168.5; DB 6; Length 1304;

```
Qy   24 PNLTEISKKITDSNAV------------LLAVKEVEALLSSSIDEIAAKAIGKKIHQNNGLDTEYN   76
Db   13 PRLSKLSASKSDSNSASPKVHSRLVKGTELQTQLNQIQEDLKKA-----------DEQ--        59

Qy   77 HNGSLLAGAYAISTLIKQK-----LDGLK------NEGLKEKIDAAKKCSETF-----      118
Db   60 IELLKDKAKAIDDLKESEKLVEEANEKLKEALAAQKRAEESFEVEKFR                 108

Qy  119 -------------TNK-----LKEKHT-DLG-----KEGVTDADAKEA                 142
Db  109 AVELEQAGLEAVQKKDVTSKNELESIRSQHALDISALLSTTEELQRVKHELSMTADAKNK     168

Qy  143 ILKTNGTKTKGAE----------ELGKL-----FESVEVLSKAAKE-----           173
Db  169 ALSHAEEATKIAEIHAEKAEILASELGRLKALLGSKEEKEAIEGNEIVSKLKSEIELLRG     228

Qy  174 ------MIANSVKE------LTSPVVAESPAMVNNSG-----KDG                    201
Db  229 ELEKVSILESSLKEQEGLVEQLKVDLEAAKMAES---CTNSSVEEKRNKVHELEKEVEES     285

Qy  202 NTSANSADESVKGPNLTEISKKITESNAVV-----LAVKEVETLLTSIDE               246
Db  286 NRSKSSASESME-----SVMKQLAELNHVLHETKSDNAAQKEKIELLEKTIEAQRTDLEE     340

Qy  247 -----LAKAIGKKIKNDV-----SLDNEADHNGSLISGAYLISNLITK                 284
Db  341 YGRQVCIAKEEASKLENLVESIKSELEISQEEKTRALDNEKAATSN----IQNLLDQ        393

Qy  285 KISAIKDSGELKAEIEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKG     344
Db  394 RTELSIELERCKVEEERSKKDMESLITLAL-----QEASTESSEAKATLLVCQEELKN       445

Qy  345 ADELEKLFESVKNLSRAAKE-----MITNSVKELTS                             375
Db  446 C---ESQVDSIKLASKETNEKYEKMLEDARNEIDS                              477
```

RESULT 11
US-09-748-875-14
Sequence 14, Application US/09748875
GENERAL INFORMATION:
; APPLICANT: BRILES et al.
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS
; TITLE OF INVENTION: AND STRAINS THEREOF AND USES THEREFOR
; FILE REFERENCE: 454312-3140
; CURRENT APPLICATION NUMBER: US/09/748,875
; CURRENT FILING DATE: 2000-12-26
; PRIOR APPLICATION NUMBER: 09/298,523
; PRIOR FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 14
; LENGTH: 589
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-748-875-14

Query Match
Best Local Similarity 8.6%; Score 158; DB 5; Length 589;

Best Local Similarity 25.1%; Pred. No. 0.0028;
Matches 109; Conservative 49; Mismatches 149; Indels 128; Gaps 20;

```
Qy   19 ESVK-GPNLT------EISKKIITDSNAVLLAVKEVEALLSSIDEIAAKAIGKKIH-QNNG 70
           |||| |||       | ||:|  :  || | |:  || | : |:  ||  :
Db    1 EGVRSGNNLIVTSSGQDISKKYAD----------EVESHLESILKDVKKNLKKVQHTQNVG 51

Qy   71 LDTE-----------------YNHNGSLLAGAYAIS----------TLIKQKLDGLKNEGLKEKIDA 110
Db   52 LITKLSEIKKKYLYDLKVNVLSEAELTSKTKETKEKLIATFEQFKKDTLPTEPEKKVAEA 111

Qy  111 AKKCSETFTNKLKEKHTD----------LGKE-GVTDADAKEAILKTNGTKTKGAEELG 158
Db  112 QKKVEEAKKKAEDQKEKDRRNYPITIYKTLELEIAESDVEVKKAELELVKKAKESQDEE 171

Qy  159 KLFES-VEVLSK------------------AAKEMLANSVKELTSPVVAES--------- 190
Db  172 KIKQAEAEVESKQAEATRLKKIKTDREEAKRKADAKLKEAVEKNVATSEQDKPKRRAKRG 231

Qy  191 ------PAMVNNSGKDGNTSANSADESVKGPNLTEISKKIITESNAVVLAVKEVETLLTS 243
Db  232 VSGELATPDKKENDAKSSDSSV--GEETLPSPSLNMANESQTEHR------KDVDEYI-- 281

Qy  244 IDELAKAIGKKIKNDVSLDNEADHNGSLISGAVLIISNLITKKISAIKDSGELKAEIEKAK 303
Db  282 ------KKMLSEIQLDRR-KHTQNV------NI-NIKLSAIKTKYLYELSVLKEN 322

Qy  304 KCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNDDKTKGADELEKLFESVKNLSKAAK 363
Db  323 SKKEELTSKTKAELT-----AAFEQFKKDTLKPE----KKVAEAEKKVEEAKKKAKDQK 372

Qy  364 E-----MLTNSVKEL 373
Db  373 EEDRRNYPTNTYKTL 387
```

RESULT 12
US-09-748-875-61
Sequence 61, Application US/09748875
GENERAL INFORMATION:
APPLICANT: BRILES et al.
TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS
TITLE OF INVENTION: AND STRAINS THEREOF AND USES THEREFOR
FILE REFERENCE: 454312-3140
CURRENT APPLICATION NUMBER: US/09/748,875
CURRENT FILING DATE: 2000-12-26
PRIOR APPLICATION NUMBER: 09/298,523
PRIOR FILING DATE: 1999-04-23
NUMBER OF SEQ ID NOS: 78
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 61
LENGTH: 690
TYPE: PRT
ORGANISM: Streptococcus pneumoniae
US-09-748-875-61

Query Match 8.6%; Score 158; DB 5; Length 690;
Best Local Similarity 25.1%; Pred. No. 0.0035;
Matches 109; Conservative 49; Mismatches 149; Indels 128; Gaps 20;

```
Qy   19 ESVK-GPNLT------EISKKIITDSNAVLLAVKEVEALLSSIDEIAAKAIGKKIH-QNNG 70
Db   38 EGVRSGNNLIVTSSGQDISKKYAD----------EVESHLESILKDVKKNLKKVQHTQNVG 88

Qy   71 LDTE-----------------YNHNGSLLAGAYAIS----------TLIKQKLDGLKNEGLKEKIDA 110
Db   89 LITKLSEIKKKYLYDLKVNVLSEAELTSKTKETKEKLIATFEQFKKDTLPTEPEKKVAEA 148

Qy  111 AKKCSETFTNKLKEKHTD----------LGKE-GVTDADAKEAILKTNGTKTKGAEELG 158
Db  149 QKKVEEAKKKAEDQKEKDRRNYPITIYKTLELEIAESDVEVKKAELELVKKAKESQDEE 208

Qy  159 KLFES-VEVLSK------------------AAKEMLANSVKELTSPVVAES--------- 190
Db  209 KIKQAEAEVESKQAEATRLKKIKTDREEAKRKADAKLKEAVEKNVATSEQDKPKRRAKRG 268

Qy  191 ------PAMVNNSGKDGNTSANSADESVKGPNLTEISKKIITESNAVVLAVKEVETLLTS 243
Db  269 VSGELATPDKKENDAKSSDSSV--GEETLPSPSLNMANESQTEHR------KDVDEYI-- 318

Qy  244 IDELAKAIGKKIKNDVSLDNEADINGSLISGAVLIISNLITKKISAIKDSGELKAEIEKAK 303
Db  319 ------KKMLSEIQLDRR-KHTQNV------NI-NIKLSAIKTKYLYELSVLKEN 359

Qy  304 KCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNDDKTKGADELEKLFESVKNLSKAAK 363
Db  360 SKKEELTSKTKAELT-----AAFEQFKKDTLKPE----KKVAEAEKKVEEAKKKAKDQK 409

Qy  364 E-----MLTNSVKEL 373
Db  410 EEDRRNYPTNTYKTL 424
```

RESULT 13
US-09-748-875-1
Sequence 1, Application US/09748875
GENERAL INFORMATION:
APPLICANT: BRILES et al.
TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS
TITLE OF INVENTION: AND STRAINS THEREOF AND USES THEREFOR
FILE REFERENCE: 454312-3140
CURRENT APPLICATION NUMBER: US/09/748,875
CURRENT FILING DATE: 2000-12-26
PRIOR APPLICATION NUMBER: 09/298,523
PRIOR FILING DATE: 1999-04-23
NUMBER OF SEQ ID NOS: 78
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 1
LENGTH: 691
TYPE: PRT
ORGANISM: Streptococcus pneumoniae
US-09-748-875-1

Query Match 8.6%; Score 158; DB 5; Length 691;
Best Local Similarity 25.1%; Pred. No. 0.0035;
Matches 109; Conservative 49; Mismatches 149; Indels 128; Gaps 20;

```
Qy  159 KLFES-VEVLSK------------------AAKEMLANSVKELTSPVVAES--------- 190
Db  209 KIKQAEAEVESKQAEATRLKKIKTDREEAKRKADAKLKEAVEKNVATSEQDKPKRRAKRG 268

Qy  191 ------PAMVNNSGKDGNTSANSADESVKGPNLTEISKKIITESNAVVLAVKEVETLLTS 243
Db  269 VSGELATPDKKENDAKSSDSSV--GEETLPSPSLNMANESQTEHR------KDVDEYI-- 318

Qy  244 IDELAKAIGKKIKNDVSLDNEADINGSLISGAVLIISNLITKKISAIKDSGELKAEIEKAK 303
Db  319 ------KKMLSEIQLDRR-KHTQNV------NI-NIKLSAIKTKYLYELSVLKEN 359

Qy  304 KCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNDDKTKGADELEKLFESVKNLSKAAK 363
Db  360 SKKEELTSKTKAELT-----AAFEQFKKDTLKPE----KKVAEAEKKVEEAKKKAKDQK 409

Qy   19 ESVK-GPNLT------EISKKIITDSNAVLLAVKEVEALLSSIDEIAAKAIGKKIH-QNNG 70
Db   38 EGVRSGNNLIVTSSGQDISKKYAD----------EVESHLESILKDVKKNLKKVQHTQNVG 88

Qy   71 LDTE-----------------YNHNGSLLAGAYAIS----------TLIKQKLDGLKNEGLKEKIDA 110
Db   89 LITKLSEIKKKYLYDLKVNVLSEAELTSKTKETKEKLIATFEQFKKDTLPTEPEKKVAEA 148

Qy  111 AKKCSETFTNKLKEKHTD----------LGKE-GVTDADAKEAILKTNGTKTKGAEELG 158
Db  149 QKKVEEAKKKAEDQKEKDRRNYPITIYKTLELEIAESDVEVKKAELELVKKAKESQDEE 208

Qy  364 E-----MLTNSVKEL 373
Db  410 EEDRRNYPTNTYKTL 424
```

RESULT 14
US-09-748-875-62
Sequence 62, Application US/09748875
GENERAL INFORMATION:
APPLICANT: BRILES et al.
TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS
TITLE OF INVENTION: AND STRAINS THEREOF AND USES THEREFOR
FILE REFERENCE: 454312-3140
CURRENT APPLICATION NUMBER: US/09/748,875
CURRENT FILING DATE: 2000-12-26
PRIOR APPLICATION NUMBER: 09/298,523
PRIOR FILING DATE: 1999-04-23
NUMBER OF SEQ ID NOS: 78
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 62
LENGTH: 701
TYPE: PRT
ORGANISM: Streptococcus pneumoniae
US-09-748-875-62

Query Match   8.6%;   Score 158;   DB 5;   Length 701;
Best Local Similarity   25.1%;   Pred. No. 0.0035;
Matches 109; Conservative 49; Mismatches 149; Indels 128; Gaps 20;

```
Qy  19 ESVK-GPNLT------EISKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGKKIH-QNNG 70
       |:| | ||||      ||||| ||          :  ||| |  | :|     ||  ||
Db  38 EGVRSGNNLTVTSSGQDISKKYAD------EVESHLESILKDVKKNLKKVQHTQNVG 88

Qy  71 LDTE--------YNHNGSLLAGAYAIS-----TLIKQKLDGLKNEGLKEKIDA 110
       |:|          :     :          :  ||:|  |  | :| ::|:   ||
Db  89 LITKLSEIKKKYLYDLKVNVLSEAELTSKTKETKEKLTATFEQFKKDTLPTEPEKKVAEA 148

Qy 111 AKKCSETFTNKLKEKHTD-------LGKE-GVTDADAKEAILKTNGTKTKGAEELG 158
       :|: :  |:|:|| ||        |  | ||||||||||||||||||||||||||
Db 149 QKKVEEAKKRAEDQKEKDRRNYPTITYKTLELEIAESDVEVKKAELELVKVKAKESQDEE 208

Qy 159 KLFES-VEVLSK---------AAKEMLANSVKELTSPVVAES--------- 190
       |:|:: ||||              ||:| || | |:  |||
Db 209 KIKQAEAEVESKQAEATRLKKIKTDREEAKRKADADAKLKEAVEKNVATSEQDKPKRRAKRG 268

Qy 191 ------PAMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLTS 243
             : ::||| :: | :|         ||||| |:|:||  ||| |: |:|:::|
Db 269 VSGELATPDKKENDAKSSDSSV--GEETLPSPSLNMANESQTEHR------KDVDEYI-- 318

Qy 244 IDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAEIEKAK 303
       | :|    :::|  ::       |     :| ||  :  :|:|| |:  | | :||
Db 319 ------KKMLSEIQLDRR-KHTQNV------NL-NIKLSAIKTKYLYELSVLKEN 359

Qy 304 KCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLSKAAK 363
       ||:|||| ||:|:|:||:|:|:||||||:|| ||:||| |||:||||||||:|||:||||
Db 360 SKKEELTSKTKAELT-----AAFEQFKKDTLKPE---KKVAEAEKKVEEAKKKAKDQK 409

Qy 364 E-----MLTNSVKEL 373
       |       |:| | |
Db 410 EEDRRNYPTNTYKTL 424
```

RESULT 15
US-09-748-875-2
Sequence 2, Application US/09748875
GENERAL INFORMATION:
APPLICANT: BRILES et al.
TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS
TITLE OF INVENTION: AND STRAINS THEREOF AND USES THEREFOR
FILE REFERENCE: 454312-3140
CURRENT APPLICATION NUMBER: US/09/748,875
CURRENT FILING DATE: 2000-12-26
PRIOR APPLICATION NUMBER: 09/298,523
PRIOR FILING DATE: 1999-04-23
NUMBER OF SEQ ID NOS: 78
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 2
LENGTH: 707
TYPE: PRT
ORGANISM: Streptococcus pneumoniae
US-09-748-875-2

Query Match   8.6%;   Score 158;   DB 5;   Length 707;
Best Local Similarity   25.1%;   Pred. No. 0.0036;
Matches 109; Conservative 49; Mismatches 149; Indels 128; Gaps 20;

```
Qy  19 ESVK-GPNLT------EISKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGKKIH-QNNG 70
       |:| | ||||      ||||| ||          :  ||| |  | :|     ||  ||
Db  38 EGVRSGNNLTVTSSGQDISKKYAD------EVESHLESILKDVKKNLKKVQHTQNVG 88

Qy  71 LDTE--------YNHNGSLLAGAYAIS-----TLIKQKLDGLKNEGLKEKIDA 110
       |:|          :     :          :  ||:|  |  | :| ::|:   ||
Db  89 LITKLSEIKKKYLYDLKVNVLSEAELTSKTKETKEKLTATFEQFKKDTLPTEPEKKVAEA 148

Qy 111 AKKCSETFTNKLKEKHTD-------LGKE-GVTDADAKEAILKTNGTKTKGAEELG 158
       :|: :  |:|:|| ||        |  | ||||||||||||||||||||||||||
Db 149 QKKVEEAKKEDQKEKDRRNYPTITYKTLELEIAESDVEVKKAELELVKVKAKESQDEE 208

Qy 159 KLFES-VEVLSK---------AAKEMLANSVKELTSPVVAES--------- 190
       |:|:: ||||              ||:| || | |:  |||
Db 209 KIKQAEAEVESKQAEATRLKKIKTDREEAKRKADADAKLKEAVEKNVATSEQDKPKRRAKRG 268

Qy 191 ------PAMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLTS 243
             : ::||| :: | :|         ||||| |:|:||  ||| |: |:|:::|
Db 269 VSGELATPDKKENDAKSSDSSV--GEETLPSPSLNMANESQTEHR------KDVDEYI-- 318

Qy 244 IDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAEIEKAK 303
       | :|    :::|  ::       |     :| ||  :  :|:|| |:  | | :||
Db 319 ------KKMLSEIQLDGR-KHTPNV------NL-NIKLSAIKTKYLYELSVLKEN 359

Qy 304 KCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLSKAAK 363
       ||:|||| ||:|:|:||:|:|:||||||:|| ||:||| |||:||||||||:|||:||||
Db 360 SKKEELTSKTKAELT-----AAFEQFKKDTLKPE---KKVAEAEKKVEEAKKKAKDQK 409

Qy 364 E-----MLTNSVKEL 373
       |       |:| | |
Db 410 EEDRRNYPTNTYKTL 424
```

Search completed: March 18, 2002, 09:58:16
Job time: 374 sec

GenCore version 4.5

Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 09:50:37 ; Search time 68.77 Seconds
                                              (without alignments)
                                              415.377 Million cell updates/sec

Title: US-09-596-746A-24
Perfect score: 1842
Sequence: 1 MACNNSGKDGNTSANSADES.........KNLSKAAKEMLTNSVKELTS 375
Scoring table:

Searched: 219241 seqs, 76174552 residues
                     Gapop 10.0 , Gapext 0.5
Scoring table: BLOSUM62

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                     Maximum Match 100%
                     Listing first 45 summaries

Database :   PIR.68:*
                     1: pir1:*
                     2: pir2:*
                     3: pir3:*
                     4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	928	50.4	210	2	G70218	outer surface prot
2	921	50.0	210	2	S69927	outer surface prot
3	879	47.7	210	2	I40144	outer surface prot
4	853	46.3	177	2	I40129	outer surface prot
5	757.5	41.1	211	2	I40145	outer surface prot
6	756.5	41.1	193	2	S70279	outer surface prot
7	696.5	37.8	193	2	S70287	outer surface prot
8	683.5	37.1	211	2	I40277	outer surface prot
9	677.5	36.8	211	2	S69918	outer surface prot
10	676.5	36.7	209	2	I40273	outer surface prot
11	674.5	36.6	191	2	S70278	outer surface prot
12	674.5	36.6	191	2	I40153	outer surface prot
13	673.5	36.6	211	2	I40268	outer surface prot
14	672.5	36.5	193	2	S70280	outer surface prot
15	666	36.2	194	2	S70277	outer surface prot
16	665	36.1	209	2	S69917	outer surface prot
17	665	36.1	190	2	S70273	outer surface prot
18	661	35.9	212	2	S70254	outer surface prot
19	661	35.9	212	2	S70276	outer surface prot
20	657.5	35.7	193	2	S70281	outer surface prot
21	655.5	35.6	211	2	I40278	outer surface prot
22	655	35.5	212	2	S69922	outer surface prot
23	653.5	35.5	191	2	S70284	outer surface prot
24	652.5	35.4	193	2	S70265	outer surface prot
25	652.5	35.4	211	2	S69932	outer surface prot
26	652	35.4	194	2	S70268	outer surface prot
27	646	35.1	212	2	I40279	outer surface prot
28	642.5	34.9	209	2	I40142	outer surface prot
29	635	34.5	212	2	I40143	outer surface prot
30	634.5	34.4	203	2	I40108	outer surface prot
31	632	34.3	214	2	S69916	outer surface prot
32	631.5	34.3	193	2	S70286	outer surface prot
33	629.5	34.2	191	2	S70288	outer surface prot
34	626	34.0	194	2	S70289	outer surface prot
35	624.5	33.9	179	2	S54222	outer surface prot
36	623.5	33.8	209	2	I40285	outer surface prot
37	621.5	33.7	209	2	I40281	outer surface prot
38	620.5	33.7	178	2	I40125	outer surface prot
39	619	33.6	210	2	S69925	outer surface prot
40	618	33.6	192	2	S70285	outer surface prot
41	617.5	33.5	178	2	I40104	outer surface prot
42	615	33.4	210	2	S69923	outer surface prot
43	614	33.3	212	2	S20543	outer surface prot
44	613	33.3	210	2	S69920	outer surface prot
45	610	33.1	210	2	I40272	outer surface prot

ALIGNMENTS

RESULT   1
G70218
outer surface protein C - Lyme disease spirochete
outer surface protein C gene sequence analysis of Borrelia burgdorferi sensu
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 21-Jul-2000
C;Accession: G70218; I40269; S37726; S70281
A;Molecule type: DNA
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Accession: I40269
A;Reference number: I40269; MUID:96025162
A;Title: Outer surface protein C gene sequence analysis of Borrelia burgdorferi sensu
J. Clin. Microbiol. 33, 2415-2420, 1995
R;Fukunaga, M.; Hanase, A.
A;Experimental source: strain B31
A;Cross-references: GB:A000792; NID:g3253098; PIDN:AAC66329.1; PID:g2689901; TIGR:BB
A;Residues: 1-210 <KLE>
A;Molecule type: DNA
A;Status: nucleic acid sequence not shown; translation not shown
A;Accession: G70218
Nature 390, 580-586, 1997
; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vu
R;Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; Wh
A;Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.
A;Reference number: A70100; MUID:98065943
A;Cross-references: GB:AE000792; NID:g3253098; PIDN:AAC66329.1; PID:g2689901; TIGR:BB
A;Residues: 1-210 <KLE>
A;Molecule type: DNA
A;Status: nucleic acid sequence not shown; translation not shown
A;Accession: S37726
A;Status: preliminary
A;Molecule type: DNA
A;Reference number: S37726; MUID:93268136
A;Title: Genetic heterogenity of the genes coding for the outer surface protein C (Os
Med. Microbiol. Immunol. 182, 37-50, 1993
R;Jauris-Heipke, S.; Fuchs, R.; Motz, M.; Preac-Mursic, V.; Schwab, E.; Soutschek, E.
A;Cross-references: GB:D49497; NID:g707092; PIDN:BAA08457.1; PID:g769684
A;Residues: 1-210 <RES>
A;Molecule type: DNA
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Accession: I40269
A;Reference number: I40269; MUID:96025162
A;Title: Outer surface protein C gene sequence analysis of Borrelia burgdorferi sensu
J. Clin. Microbiol. 33, 2415-2420, 1995
R;Livey, I.; Gibbs, C.P.; Schuster, R.; Dorner, F.
A;Cross-references: EMBL:X69596; NID:g311391; PIDN:CAA49306.1; PID:g311392
A;Residues: 1-210 <JAU>
A;Molecule type: DNA
A;Status: preliminary
A;Accession: S37726
A;Reference number: S70255; MUID:96296448
A;Title: Evidence for lateral transfer and recombination in OspC variation in Lyme di
Mol. Microbiol. 18, 257-269, 1995
R;Livey, I.; Gibbs, C.P.; Schuster, R.; Dorner, F.
A;Cross-references: EMBL:L42887; NID:g858715; PIDN:AAB36995.1; PID:g1695212
A;Residues: 19-210 <LIV>
A;Status: nucleic acid sequence not shown
A;Accession: S70281
A;Molecule type: DNA
A;Residues: 19-210 <LIV>
A;Cross-references: EMBL:L42887; NID:g858715; PIDN:AAB36995.1; PID:g1695212
A;Experimental source: strain Ip2
C;Genetics:
A;Gene: ospC
C;Superfamily: Lyme disease spirochete surface protein C

RESULT 3

Qy 1 MACNNSGKDGNTSANSADESVKGPNLTEISKKIITDSNAVILLAVKEVEALLSSIDEIAAKA
      :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 17 ISCNNSGKDGNTSANSADESVKGPNLTEISKKIITDSNAVILLAVKEVEALLSSIDEIAAKA 60

Qy 61 IGKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN 120
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 77 IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN 136

Qy 121 KLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 180
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 137 KLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 196

Qy 181 ELTSPVVAESP 191
      |||||||||||
Db 197 ELTSPVVAESP 207

Query Match   50.0%; Score 921; DB 2; Length 210;
Best Local Similarity   97.4%; Pred. No. 3.6e-36;
Matches 186; Conservative   4; Mismatches   1; Indels   0; Gaps   0;

RESULT 2

S69927

outer surface protein C precursor - Lyme disease spirochete (strain PKa)
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
A;Variety: strain Pka
C;Date: 15-Feb-1997 #sequence_revision 27-Feb-1997 #text_change 26-May-2000
C;Accession: S69927; S72669
R;Jauris-Heipke, S.; Liegl, G.; Preac-Mursic, V.; Roessler, D.; Schwab, E.; Soutschek,
J. Clin. Microbiol. 33, 1860-1866, 1995
A;Title: Molecular analysis of genes encoding outer surface protein C (OspC) of Borrelia
A;Reference number: I40047; MUID:9539J5018
A;Accession: S69927
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-210 <JAU>
A;Cross-references: EMBL:X69589
A;Experimental source: strain Pka
R;Jauris, S.
submitted to the EMBL Data Library, February 1994
A;Reference number: S72669
A;Accession: S72669
A;Molecule type: DNA
A;Residues: 1-124,'D',126-139,'E',141-210 <JAW>
A;Cross-references: EMBL:X69589
C;Genetics:
ne: ospC
perfamily: Lyme disease spirochete surface protein C

outer surface protein C precursor - Lyme disease spirochete
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 21-Jul-2000
C;Accession: I40144; S70282
R;Stevenson, B.; Barthold, S. W.
FEMS Microbiol. Lett. 124, 367-372, 1994
A;Title: Expression and sequence of outer surface protein C among North American isol
A;Reference number: I40143; MUID:95154673
A;Accession: I40144
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-210 <RES>
A;Cross-references: EMBL:U04281; NID:g434663; PIDN:AAC43297.1; PID:g434664
Mol. Microbiol. 18, 257-269, 1995
R;Livey, I.; Gibbs, C.P.; Schuster, R.; Dorner, F.
A;Title: Evidence for lateral transfer and recombination in OspC variation in Lyme di
A;Reference number: S70255; MUID:96296448
A;Accession: S70282
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 19-210 <LIV>
A;Cross-references: EMBL:L42893; NID:g858721; PIDN:AAB37001.1; PID:g1695218
A;Experimental source: strain 297
C;Genetics:
A;Gene: ospC
C;Superfamily: Lyme disease spirochete surface protein C

Qy 195 NNSGKDGNTSANSADESVKGPNLTEISKKIITESNAVVLAVKEVETLLTSIDELAKAIGKK 254
        ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db 20 NNSGKDGNTSANSADESVKGPNLTEISKKIITESNAVVLAVKEVETLLTSIDELAKAIGKK 79

Qy 255 IKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAEIEKAKKCSEEFTARLK 314
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 80 IKNDVSLDNEADHNGSLISGAYLISTLITKKISAIKDSGELKAEIEKAKKCSEEFTARLK 139

Qy 315 GEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLSKAAKEMLTNSVKELT 374
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 140 GEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLSKAAKEMLTNSVKELT 199

Qy 375 S 375
        |
Db 200 S 200

Query Match   47.7%; Score 879; DB 2; Length 210;
Best Local Similarity   99.4%; Pred. No. 2.8e-34;
Matches 180; Conservative   0; Mismatches   1; Indels   0; Gaps   0;

RESULT 4

I40129

outer surface protein C - Lyme disease spirochete (strain TXGW) (fragment)
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 26-May-2000
C;Accession: I40129; S54199
R;Theisen, M.; Borre, M.; Mathiesen, M.J.; Mikkelsen, B.; Lebech, A.M.; Hansen, K.
J. Bacteriol. 177, 3036-3044, 1995
A;Title: Evolution of the Borrelia burgdorferi outer surface protein OspC.
A;Reference number: I40104; MUID:95286481
A;Accession: I40129
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-177 <RES>
A;Cross-references: EMBL:X84783; NID:g793825; PIDN:CAA59254.1; PID:g793826
C;Genetics:
A;Gene: ospC
C;Superfamily: Lyme disease spirochete surface protein C

Qy 1 MACNNSGKDGNTSANSADESVKGPNLTEISKKIITDSNAVILLAVKEVEALLSSIDEIAAKA
      :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 17 ISCNNSGKDGNTSANSADESVKGPNLTEISKKIITDSNAVILLAVKEVEALLSSIDEIAAKA 60

Qy 61 IGKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIIDAAKKCSETFTN 120
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 77 IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIEAAKKCSETFTN 136

Qy 121 KLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 180
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 137 KLKDKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 180

Qy 181 ELTSPVVAESP 191
      |||||||||||
Db 197 ELTSPVVAESP 207

Query Match   46.3%; Score 853; DB 2; Length 177;
Best Local Similarity   98.9%; Pred. No. 3.6e-33;
Matches 175; Conservative   0; Mismatches   2; Indels   0; Gaps   0;

RESULT

I40144

```
Qy    11 NTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGKKIHQNNG 70
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 NTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGKKIHQNNG 60

Qy    71 LDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSEFFTNKLKEKHTDLG 130
         |||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 LDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSEFFTNKLKEKHTDLG 120

Qy   131 KEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKELTSPVV 187
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 KRGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKELTSPVV 177
```

RESULT 5

I40145
outer surface protein C precursor - Lyme disease spirochete
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 21-Jul-2000
C;Accession: I40145
C;Genetics:
A;Gene: ospC
C;Superfamily: Lyme disease spirochete surface protein C

A;Accession: I40145
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-211 <RES>
A;Cross-references: EMBL:U04282; NID:g2314881; PIDN:AAC45540.1; PID:g434666
R;Title: Expression and sequence of outer surface protein C among North American isolate
A;Reference number: I40143; MUID:9514673
R;Stevenson, B.; Barthold, S.W.
R;Microbiol. Lett. 124, 367-372, 1994

Query Match   41.1%; Score 757.5; DB 2; Length 211;
Best Local Similarity   81.2%;
Matches 156; Conservative 13; Mismatches 22; Indels   1; Gaps   1;

Pred. No. 1.1e-28;

RESULT 6

S70279
outer surface protein C - Lyme disease spirochete
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 26-May-2000
C;Accession: S70279
C;Genetics:
A;Gene: ospC
A;Accession: S70279
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-193 <LIV>
A;Cross-references: EMBL:L42898; NID:g858729; PIDN:AAB37007.1; PID:g1695223
A;Experimental source: strain 25015
A;Reference number: S70255; MUID:96296448
A;Title: Evidence for lateral transfer and recombination in OspC variation in Lyme disea
R;Livey, I.; Gibbs, C.P.; Schuster, R.; Dorner, F.
R;Mol. Microbiol. 18, 257-269, 1995

Query Match   41.1%; Score 756.5; DB 2; Length 193;
Best Local Similarity   82.1%; Pred. No. 1.1e-28;
Matches 156; Conservative 10; Mismatches 23; Indels   1; Gaps   1;

```
Qy     1 MACNNSGKDGN-TSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAK 59
          :||||||||| ||||| ||||||||||||||||||||: ||||||| ||||:||||:| |
Db    17 ISCNNSGKDGNAASTNPADESVKGPNLTEISKKITDSNTVVLAVKEVGALLTSIDELATK 76

Qy    60 AIGKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSEFFT 119
         |||||||||||||| ||||||||||||||||||:||| |||||:|||||:|| ||||||
Db    77 AIGKKIHQNNGLDTENNHNGSLLAGAYAISTLITQKLGGLKNEELKEKIAAVKKCSEFFT 136

Qy   120 NKLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSV 179
         ||||.. |.||:. ::.|||:||| :|:||| .||||| |.||:||:| |||||| :||||
Db   137 NKLKSSHTELGKQDAQDDDAKKAILRTHNTKDKGAEELDKLFKPVENLSKAAKEMLSNSV 196

Qy   180 KELTSPVVAESP 191
         ||||||||||||
Db   197 KELTSPVVAESP 208
```

RESULT 7

S70287
outer surface protein C - Lyme disease spirochete
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 26-May-2000
C;Accession: S70287
C;Genetics:
A;Gene: ospC
A;Accession: S70287
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-193 <LIV>
A;Cross-references: EMBL:L42895; NID:g858723; PIDN:AAB37003.1; PID:g1695220
A;Experimental source: strain 28354
A;Reference number: S70255; MUID:96296448
R;Livey, I.; Gibbs, C.P.; Schuster, R.; Dorner, F.
R;Mol. Microbiol. 18, 257-269, 1995
C;Superfamily: Lyme disease spirochete surface protein C

Query Match   41.1%; Score 756.5; DB 2; Length 193;
Best Local Similarity   82.1%; Pred. No. 1.1e-28;
Matches 156; Conservative 10; Mismatches 23; Indels   1; Gaps   1;

```
Qy     3 CNNSGKDGN-TSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAK 59
         |||||||| ||||| ||||||||||||||||||||: ||||||| ||||:||||:|
Db     1 CNNSGKDGNAASTNPADESVKGPNLTEISKKITDSNTVVLAVKEVGALLTSIDELATK 76

Qy    60 AIGKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSEFFT 119
         |||||||||||||| ||||||||||||||||||:||| |||||:|||||:|| ||||||
Db    77 AIGKKIHQNNGLDTENNHNGSLLAGAYAISTLITQKLGGLKNEELKEKIAAVKKCSEFFT 136

Qy   120 NKLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSV 179
         ||||.. |.||:. ::.|||:||| :|:||| .||||| |.||:||:| |||||| :||||
Db   137 NKLKSSHTELGKQDAQDDDAKKAILRTHNTKDKGAEELDKLFKPVENLSKAAKEMLSNSV 196

Qy   180 KELTSPVVAESP 191
         ||||||||||||
Db   181 KELTSPVVAESP 190
```

RESULT 8

I40277
outer surface protein C precursor - Borrelia afzelii
C;Species: Borrelia afzelii
C;Date: 04-Sep-1997 #sequence_revision 04-Sep-1997 #text_change 26-May-2000
C;Accession: I40277

Query Match   37.8%; Score 696.5; DB 2; Length 193;
Best Local Similarity   76.3%; Pred. No. 6.4e-26;
Matches 145; Conservative 16; Mismatches 28; Indels   1; Gaps   1;

```
Qy     3 CNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAIG 62
         |||||||||||||||||||||||||||||||||:||||:||||:|  |::|||||:||||
Db     1 CNNSGKDGNTSANSADESVKGPNLTEISKKITPESNAVVLAVKEIENAKAKAAKEMLANSVKE 60

Qy    63 KKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKN-EGLKEKIDAAKKCSEFFTNK 121
         |||.|| | : :.:||:.|||||||:|: ||  : : |  :||||||:.||||||| ||
Db    61 KKIQQNGGLAVEAGHNGTLLAGAYTISKLITGAYTISKLITDAAKDGAAELEKLFKAVENLAKAAKEMLANSVKE 120

Qy   122 LKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 181
         |.|:| |||.|:||| :||:|||.|: | .|||| |:|||:.|||| ||||||||||||||
Db   121 LEGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLAKAAKEMLANSVKE 180

Qy   182 LTSPVVAESP 191
         ||||:|||||
Db   181 LTSPIVAESP 190
```

A;Title: Outer surface protein C gene sequence analysis of Borrelia burgdorferi sensu la
A;Reference number: I40269; MUID:96025162
A;Accession: I40277
A;Molecule type: translated from GB/EMBL/DDBJ
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-211 <RES>
A;Cross-references: GB:D49501; NID:g7070096; PIDN:BAA08461.1; PID:g769688
C;Superfamily: Lyme disease spirochete surface protein C

```
RESULT 9
S69918

Query Match   37.1%; Score 683.5; DB 2; Length 211;
Best Local Similarity 75.1%; Pred. No. 2.8e-25;
Matches 145; Conservative 19; Mismatches 26; Indels 3; Gaps 3;

Qy   1 MACNNSGKDGNT-SANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAK 59
         :|||| || |: :||||||||||||||||||||||||||||:|||||||||||| |
Db  17 ISCNNSGKGGDSASTNPDESARGPNLTEISKKITDSNAFVLAVKEVETLVSSIDELATK 76

Qy  60 AIGKKIHQNNGLDTEYNINGSLLAGAYAISTLIKQKLDGLKN-EGLKEKIDAAKKCSETF 118
         |||||:   |  ||  |||||||||||||||||||||||||  |||||||||||||||
Db  77 AIGKI-ONNGLAAAEADKNGSLLAGAYAISTLIKQKLDGLKGLEGLNKEIAEAKKCSEAF 135

Qy 119 TNKLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANS 178
         |||||:|||||||||:|||:||| : ||  ||| |||||| | ||| |||||||||||
Db 136 TKLKQDSNADLGKHNATDADSKEAILKTNGTKTKGAKELEELFKSVESLSKAAKEALSNS 195

Qy 179 VKELTSPVVAESP 191
         |||||||||||||
Db 196 VKELTSPVVAESP 208
```

outer surface protein C precursor - Lyme disease spirochete (strain PBre)
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
A;Variety: strain PBre
C;Date: 06-Dec-1996 #sequence_revision 14-Feb-1997 #text_change 26-May-2000
R;Jauris-Heipke, S.; Liegl, G.; Preac-Mursic, V.; Roessler, D.; Schwab, E.; Soutschek, E
J. Clin. Microbiol. 33, 1860-1866, 1995
A;Accession: S69918; S72674; I40103
A;Reference number: I40047; MUID:95395018
A;Title: Molecular analysis of genes encoding outer surface protein C (OspC) of Borrelia
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-211 <JAU>
A;Cross-references: EMBL:X81522; NID:g872021; PIDN:CAA57242.1; PID:g872022

Roessler, D.
submitted to the EMBL Data Library, September 1994
A;Reference number: S72674
A;Molecule type: DNA
A;Residues: 1-152, 'E', 154-211 <ROE>
A;Cross-references: EMBL:X81522; NID:g872021; PIDN:CAA57242.1; PID:g872022

C;Superfamily: Lyme disease spirochete surface protein C

```
RESULT 10
I40273

Query Match   36.7%; Score 676.5; DB 2; Length 209;
Best Local Similarity 75.0%; Pred. No. 5.9e-25;
Matches 144; Conservative 20; Mismatches 25; Indels 3; Gaps 3;

Qy   1 MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA 60
         :||||  |||  :||||||||||||||||||||||||||||:|||||||||||| ||
Db  17 ISCNNSGKGGDIASTNPDESAKGPNLTEISKKITDSNAVVLAVKEVETLVSSIDEL-AKT 75

Qy  61 IGKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKN-EGLKEKIDAAKKCSETFT 119
         |||: |  ||  ||  ||||||||||||||||||||||||  ||||||||||||||||
Db  76 IGKKI-EANGLGNEADKNGSLLAGAYAISTLIKQKLDGLKGLEGLNKEIAEAKKCSEAFT 134

Qy 120 NKLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSV 179
         |||:|:|||||||||:|||:|||| | |||||| |||||| | ||| ||||||||||||
Db 135 KKLQDSNADLGKHNATDADSKEAILKTNGTKTKGAKELEELFKSVESLSKAAKEALSNSV 194

Qy 180 KELTSPVVAESP 191
         ||||||||||||
Db 195 KELTSPVVAESP 206
```

outer surface protein C precursor - Borrelia afzelii
C;Species: Borrelia afzelii
C;Date: 04-Sep-1997 #sequence_revision 04-Sep-1997 #text_change 26-May-2000
R;Fukunaga, M.; Hamase, A.
J. Clin. Microbiol. 33, 2415-2420, 1995
A;Title: Outer surface protein C gene sequence analysis of Borrelia burgdorferi sensu
A;Reference number: I40269; MUID:96025162
A;Accession: I40273
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-209 <RES>
A;Cross-references: GB:D49379; NID:g1041107; PIDN:BAA08377.1; PID:g1041108
C;Superfamily: Lyme disease spirochete surface protein C

```
RESULT 11
S70278

Query Match   36.6%; Score 674.5; DB 2; Length 191;
Best Local Similarity 75.8%; Pred. No. 6.5e-25;
Matches 144; Conservative 18; Mismatches 25; Indels 3; Gaps 3;

Qy   1 MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA 60
         :||||| || |:  :||||||||||||||||||||||||||:|||||||||||| || |
Db  17 ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAIG 62

Qy  61 IGKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQRLDGLIK -NEGLKEKIDAAKKCSETFT 119
         |||| |  ||  ||  |||||||||||||:| ||||| :   |||| || || || |
Db  76 IGKKIKNDGSLGDEANHNESLLAGAYTISTLIQQKLSKLNGSEGLKEKIEAALSSIDEL-AKA 75

Qy 120 NKLKEKHTDLGKEGVTDADADAKEAILKTNGT-KTKGAEELGKLFESVEVLSKAAKEMLANS 178
         |||:|:||||||:||::    | ||||||| ||||||| || | | || || || || || |
Db 135 IGKKIKNDGSLGDEANHNESLLAGAYTISTLIQQKLSKLNGSEGLKEKIEAALSSIDEL-AKA 135
```

outer surface protein C - Lyme disease spirochete
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
A;Variety: strain VS461
C;Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 26-May-2000
R;Livey, I.; Gibbs, C.P.; Schuster, R.; Dorner, F.
Mol. Microbiol. 18, 257-269, 1995
A;Title: Evidence for lateral transfer and recombination in OspC variation in Lyme di
A;Reference number: S70278
A;Accession: S70255; MUID:96296448
A;Accession: S70278
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-191 <LIV>
A;Experimental source: strain VS461
A;Cross-references: EMBL:L42871; NID:g858738; PIDN:AAB37014.1; PID:g1695229
C;Gene: ospC
C;Genetics:
C;Superfamily: Lyme disease spirochete surface protein C

RESULT 12

I40153

outer surface protein C - Lyme disease spirochete (fragment)
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-191 <RES>
C;Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 26-May-2000
Accession: I40153
C;Cross-references: EMBL:U08284; NID:g469561; PIDN:AAA21460.1; PID:g469562
A;Reference number: I40153; MUID:94314437
A;Title: Humoral immune response to outer surface protein C of Borrelia burgdorferi in l
Infect. Immun. 62, 3213-3221, 1994
Ung, B.P.; McHugh, G.L.; Leong, J.M.; Steere, A.C.
A;Genetics:
A;Gene: OspC
C;Superfamily: Lyme disease spirochete surface protein C

```
                                              Best Local Similarity   36.6%;        Score 674.5;  DB 2;     Length 191;
                                                       75.5%;   Pred. No. 6.5e-25;
                                    Matches 142;  Conservative  16;  Mismatches  29;  Indels   1;   Gaps   1;

Qy      5  NSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGKK  64
           |||||||||||||||||||||||||||||||||||||||||||||:|:|||||||||||||
Db      1  NSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDELATKAIGKK  60

Qy     65  IHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKN-EGLKEKIDAAKKCSETFTNKLK  123
           ||||||||||||:||||||||||||||||||||||||  |||:|||||||||||||||:||
Db     61  IQQNGGLAVEAGHNGTLLAGAVTISKLITQKLDGLKNSEKLKEKIENAKKCSEDFTKKLE  120

Qy    124  EKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKELT  183
           ::||||:|| ||:|:|||||:||||:||||||||||||||||||||||||||||||||||
Db    121  GEHAQLGIENVTDENAKRAILITDAAKDKGAAELEKLFKAVENLAKAAKEMLANSVKELT  180

Qy    184  SPVVAESP  191
           ||:|||||
Db    181  SPIVAESP  188
```

RESULT 13

I40268

outer surface protein C precursor - Lyme disease spirochete
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
A;Accession: I40268
R;Margolis, N.; Hogan, D.; Cieplak, W.
Gene 143, 105-110, 1994
A;Title: Homology between Borrelia burgdorferi OspC and members of the family of Borreli
A;Reference number: I40268; MUID:94259285
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-211 <RES>
C;Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 26-May-2000
C;Accession: I40268
C;Cross-references: GB:L25413; NID:g495735; PIDN:AAA22956.1; PID:g495736
C;Genetics:

RESULT 14

S70280

outer surface protein C - Lyme disease spirochete
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 26-May-2000
C;Accession: S70280
A;Genetics:
A;Gene: OspC
C;Superfamily: Lyme disease spirochete surface protein C

A;Gene: OspC
C;Superfamily: Lyme disease spirochete surface protein C

```
                                              Best Local Similarity   36.6%;        Score 673.5;  DB 2;     Length 211;
                                                       74.5%;   Pred. No. 8.1e-25;
                                    Matches 143;  Conservative  16;  Mismatches  32;  Indels   1;   Gaps   1;

Qy      1  MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA  60
           :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|
Db     17  ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEVLLSDELAKKA  76

Qy     61  IGKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEG-LKEKIDAAKKCSETFT  119
           |||||:|||||:|||||||||||||||||||||||||||||||  ||||||||||||||:
Db     77  IGKKIDQNNALGTLDNHNGSLLAGAVAISALITEKLSSIKDSGELKAEIERAKKCSESFT  136

Qy    120  NKLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSV  179
           ||:|  ::|||||||||:|||||||:||||:||||||||||||||||||||||||||||
Db    137  KKLSDNQAELGIENATDDNAKKAILKTHNAKDKGAEELVKLSESVAGLLKAAQAILANSV  196

Qy    180  KELTSPVVAESP  191
           |||||:||||||
Db    197  KELTSPVVAESP  208
```

A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-193 <LIV>
A;Accession: S70280
A;Reference number: S70255; MUID:96296448
A;Title: Evidence for lateral transfer and recombination in OspC variation in Lyme di
Mol. Microbiol. 18, 257-269, 1995
R;Livey, I.; Gibbs, C.P.; Schuster, R.; Dorner, F.
A;Cross-references: EMBL:L42868; NID:g858735; PIDN:AAB37011.1; PID:g1695226
A;Genetics:
A;Experimental source: strain ZS7
A;Cross-references:

RESULT 15

S70277

outer surface protein C - Lyme disease spirochete
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 26-May-2000
C;Accession: S70277

```
                                              Best Local Similarity   36.5%;        Score 672.5;  DB 2;     Length 193;
                                                       77.5%;   Pred. No. 8.1e-25;
                                    Matches 148;  Conservative   9;  Mismatches  31;  Indels   3;   Gaps   3;

Qy      3  CNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAIG  62
           |||||||||||||||||||||||||||||||||||||||:|||||||||||||||||:||
Db      1  CNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEVLLSSGELSESLSKAAKEMLANSVK  59

Qy     63  KKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLK-NEGLKEKIDAAKKCSEFSTK  121
           |||::|||:||||||||||||||||||||||:|||| |  ||||||||||||||||:||
Db     60  KKIKNDGSLGDEANINESLLAGAYTISTLITPQKLSKLNGSEGLKEKIAAAKKCSEFSTK  119

Qy    122  LKEKHTDLGKEGVTDADAKEAILKTNGT-KTKGAEELGKLFESVEVLSKAAKEMLANSVK  180
           ||:|||||||||||:|:|||||||||||  ||||||||||||||||||||||||||||||
Db    120  LKDNHAQLGIGVTDENAKRAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANSVK  179

Qy    181  ELTSPVVAESP  191
           ||||||||:|||
Db    180  ELTSPVVESP  190
```

R;Livey, I.; Gibbs, C.P.; Schuster, R.; Dorner, F.
Mol. Microbiol. 18, 257-269, 1995
A;Title: Evidence for lateral transfer and recombination in OspC variation in Lyme disea
A;Reference number: S70255; MUID:9629644 8
A;Accession: S70277
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-194 <LIV>
A;Cross-references: EMBL:L42873; NID:g858740; PIDN:AAB37016.1; PID:g1695231
A;Experimental source: strain SIMON
C;Genetics:
A;Gene: ospC
C;Superfamily: Lyme disease spirochete surface protein C

Query Match   36.2%; Score 666; DB 2;   Length 194;
Best Local Similarity   74.3%; Pred. No. 1.6e-24;
Matches 142; Conservative 13; Mismatches 34; Indels 2; Gaps 2;

```
Qy    3 CNNSGKDG-NTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAI 61
         ||||| | :||| | ||||||||||||||| :|||||||| |::||||||:|:|| |||
Db    1 CNNSGKGGDSTSTNPADESAKGPNLTEISKKITNSNAFVLAVKEVETLVASIDELATKAI 60

Qy   62 GKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKN-EGLKEKIDAAKKCSETFTN 120
         ||||   : |||||||||||||||||||| ||| ::| |||||||||||||||
Db   61 GKKIKNDGTLENEANHNGSLLAGAYAISNLIKQKLDGSLKGLEGLINKEIAEAKNCSEAFTK 120

Qy  121 KLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 180
         |||||||||||| |||| | |||| | |||| | ||| |||| | |||| |||:-| |||
Db  121 KLKEKHTDLGKENATDEDAKKAILKTDATKDKGAAELEKLSESVASILVKAAQEALTNSVK 180

Qy  181 ELTSPVVAESP 191
         |||||||||||
Db  181 ELTSPVVAESP 191
```

Search completed: March 18, 2002, 09:56:46
Job time: 369 sec

OM protein - protein search, using sw model

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

Run on: March 18, 2002, 09:54:29 ; Search time 118.14 Seconds
                                          (without alignments)
                                          237.004 Million cell updates/sec

Title: 1859
Perfect score: 1859
Sequence: 1 MACNNSGKDGNTSANSADES.........KNLSKAAKEMLTNSVKELTS 378

Scoring table: BLOSUM62
               Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   A_Geneseq_1101:*
         1: /SIDS8/gcgdata/geneseq/geneseqp/AA1980.DAT:*
         2: /SIDS8/gcgdata/geneseq/geneseqp/AA1981.DAT:*
         3: /SIDS8/gcgdata/geneseq/geneseqp/AA1982.DAT:*
         4: /SIDS8/gcgdata/geneseq/geneseqp/AA1983.DAT:*
         5: /SIDS8/gcgdata/geneseq/geneseqp/AA1984.DAT:*
         6: /SIDS8/gcgdata/geneseq/geneseqp/AA1985.DAT:*
         7: /SIDS8/gcgdata/geneseq/geneseqp/AA1986.DAT:*
         8: /SIDS8/gcgdata/geneseq/geneseqp/AA1987.DAT:*
         9: /SIDS8/gcgdata/geneseq/geneseqp/AA1988.DAT:*
        10: /SIDS8/gcgdata/geneseq/geneseqp/AA1989.DAT:*
        11: /SIDS8/gcgdata/geneseq/geneseqp/AA1990.DAT:*
        12: /SIDS8/gcgdata/geneseq/geneseqp/AA1991.DAT:*
        13: /SIDS8/gcgdata/geneseq/geneseqp/AA1992.DAT:*
        14: /SIDS8/gcgdata/geneseq/geneseqp/AA1993.DAT:*
        15: /SIDS8/gcgdata/geneseq/geneseqp/AA1994.DAT:*
        16: /SIDS8/gcgdata/geneseq/geneseqp/AA1995.DAT:*
        17: /SIDS8/gcgdata/geneseq/geneseqp/AA1996.DAT:*
        18: /SIDS8/gcgdata/geneseq/geneseqp/AA1997.DAT:*
        19: /SIDS8/gcgdata/geneseq/geneseqp/AA1998.DAT:*
        20: /SIDS8/gcgdata/geneseq/geneseqp/AA1999.DAT:*
        21: /SIDS8/gcgdata/geneseq/geneseqp/AA2000.DAT:*
        22: /SIDS8/gcgdata/geneseq/geneseqp/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB ID	Description
1	1859	100.0	378	22	AAB62712	Borrelia sp chimer
2	1834	98.7	401	22	AAB62738	Borrelia sp chimer
3	1805.5	97.1	374	22	AAB62710	Borrelia sp chimer
4	1802.5	97.0	398	22	AAB62728	Borrelia sp chimer
5	1614.5	86.8	377	22	AAB62713	Borrelia sp chimer
6	1589.5	85.5	400	22	AAB62739	Borrelia sp chimer
7	1561	84.0	373	22	AAB62711	Borrelia sp chimer
8	1558	83.8	397	22	AAB62729	Borrelia sp chimer
9	1528.5	82.2	370	22	AAB62715	Borrelia sp chimer
10	1525.5	82.1	394	22	AAB62730	Borrelia sp chimer
11	1496.5	80.5	384	22	AAB62726	Borrelia sp chimer
12	1493.5	80.3	408	22	AAB62737	Borrelia sp chimer
13	1489.5	80.1	368	22	AAB62718	Borrelia sp chimer
14	1486.5	80.0	392	22	AAB62734	Borrelia sp chimer
15	1460.5	78.6	560	22	AAB62724	Borrelia sp chimer
16	1284	69.1	369	22	AAB62716	Borrelia sp chimer
17	1281	68.9	393	22	AAB62731	Borrelia sp chimer
18	1245	67.0	367	22	AAB62719	Borrelia sp chimer
19	1242	66.8	391	22	AAB62735	Borrelia sp chimer
20	1213	65.3	378	22	AAB62714	Borrelia sp chimer
21	1209	65.2	393	22	AAB62725	Borrelia sp chimer
22	1205	64.8	401	22	AAB62733	Borrelia sp chimer
23	1193.5	64.2	410	22	AAB62740	Borrelia sp chimer
24	1192.5	64.1	386	22	AAB62727	Borrelia sp chimer
25	1178	63.4	368	22	AAB62717	Borrelia sp chimer
26	1170	62.9	391	22	AAB62736	Borrelia sp chimer
27	968.5	52.1	192	16	AAB62721	Borrelia IP2 OspC
28	954	51.3	209	16	AAB62720	Outer surface prot
29	945	50.8	210	16	AAB62736	B burgdorferi OspC
30	943	50.7	192	15	AAW11935	B31 outer surface
31	939	50.5	192	15	AAR60886	B31 outer surface
32	929	50.0	210	16	AAB62721	B. burgdorferi OspC
33	928	49.9	210	16	AAR75746	B31 OspC
34	915	49.2	466	16	AAR75727	B31 Osp-A/antigen
35	908	48.8	192	22	AAR75739	B. burgdorferi str
36	885	47.6	190	22	AAB62703	B burgdorferi OspC
37	879	47.3	192	15	AAB62707	B. burgdorferi OspC
38	864	46.5	192	15	AAR60889	Borrelia 297 antig
39	852	45.8	176	15	AAR62772	Borrelia B31 antig
40	694.5	37.4	176	15	AAR62775	Borrelia 297 antig
41	694.5	37.4	177	15	AAR62773	Borrelia 25015 ant
42	692.5	37.3	177	15	AAR60887	Borrelia 25015 Osp
43	692.5	37.3	211	18	AAW41822	B. burgdorferi sen
44	690.5	37.1	211	20	AAW93473	B. burgdorferi Osp
45	690.5	37.1	193	15	AAR60888	Borrelia 2S7 antig

ALIGNMENTS

RESULT 1
AAB62712

ID   AAB62712 standard; Protein; 378 AA.
XX
AC   AAB62712;
XX
DT   03-APR-2001 (first entry)
XX
DE   Borrelia sp chimeric ospC protein SEQ ID NO: 28.
KW   Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
OS   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
XX
PN   WO200078966-A1.
XX
PD   28-DEC-2000.
XX
PF   19-JUN-2000; 2000WO-US16915.
XX
PR   18-JUN-1999; 99US-0140042.
XX
PA   (UYNY) UNIV NEW YORK STATE RES FOUND.
PA   (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI   Dattwyler RJ, Seinost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
DR   WPI; 2001-050113/06.
DR   N-PSDB; AAF29016.
XX
PT   Compositions of OspC polypeptides from strains of Borrelia which cause
PT   Lyme disease are used to immunize animals and detect immune responses
PT   to Lyme disease -

RESULT 2

XX
PS Claim 43; Page 83-84; 160pp; English.
XX
CC The present invention provides compositions comprising ospc proteins and
CC chimeric ospc proteins from members of the Borrelia genus. These may be
CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC vaccines against Borrelia infection, which is spread by ticks and leads
CC to Lyme disease.
XX
SQ Sequence 378 AA;

Query Match
Best Local Similarity 100.0%;   Score 1859; DB 22; Length 378;
Matches 378; Conservative 0; Pred. No. 2.9e-121;
Mismatches 0; Indels 0; Gaps 0;

```
Qy    1 MACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKA 60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 macnnsgkdgntsansadesvkgpnlteinkkitdsnavllavkeveallssideiaaka 60

Qy   61 IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN 120
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 igkkihqnnglidtennhngsllagayaistlikqkldglkneglkekidaakkcsetftn 120

Qy  121 KLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 klkekhtdlgkegvtdadakeailkangtktkgaeelgklfesvevlskaakemlansvk 180

Qy  181 ELTSPVVAESPKKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEV 240
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 eltspvvaespkkpsmvnnsgkdgntsansadesvkgpnlteiskkitesnavvlavkev 240

Qy  241 ETLLTSIDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKA 300
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 etlltsidelakaigkkikndvslidneadhngslisgayliisnlitkkisaikdsgelka 300

Qy  301 EIEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKN 360
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 eiekakkcseeftaklkgehtdlgkegvtddnakkailktnndktkgadeleklfesvkn 360

Qy  361 LSKAAKEMLTNSVKELTS 378
        ||||||||||||||||||
Db  361 lskaakemltnsvkelts 378
```

RESULT 2
AAB62738
AAB62738 standard; Protein; 401 AA.
AC AAB62738;
XX
DT 03-APR-2001 (first entry)
XX
DE Borrelia sp chimeric ospc protein SEQ ID NO: 80.
KW Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
OS Chimeric - Borrelia sp.
OS Chimeric - Borrelia sp.
XX
PN WO200078966-A1.
XX
PD 28-DEC-2000.
XX
PF 19-JUN-2000; 2000WO-US16915.
XX
PR 18-JUN-1999; 99US-0140042.
XX
PA (UYNY) UNIV NEW YORK STATE RES FOUND.
PA (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI Dattwyler RJ, Selnost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;

RESULT 3

DR WPI; 2001-050113/06.
DR N-PSDB; AAF29042.
XX
PT Compositions of OspC polypeptides from strains of Borrelia which cause
PT Lyme disease are used to immunize animals and detect immune responses
XX
PS Claim 43; Page 150-151; 160pp; English.
XX
CC The present invention provides compositions comprising ospc proteins and
CC chimeric ospc proteins from members of the Borrelia genus. These may be
CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC vaccines against Borrelia infection, which is spread by ticks and leads
CC to Lyme disease.
XX
SQ Sequence 401 AA;

Query Match
Best Local Similarity 98.7%;   Score 1834; DB 22; Length 401;
Matches 373; Conservative 2; Pred. No. 1.7e-119;
Mismatches 2; Indels 2; Gaps 0;

```
Qy    2 ACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKAI 61
        :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   25 scnnsgkdgntsansadesvkgpnlteinkkitdsnavllavkeveallssideiaakai 84

Qy   62 GKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNK 121
        ||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db   85 gkkihqnnglidteynhngsllagayaistlikqkldglkneglkekidaakkcsetftnk 144

Qy  122 LKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 181
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  145 lkekhtdlgkegvtdadakeailkangtktkgaeelgklfesvevlskaakemlansvke 204

Qy  182 LTSPVVAESPKKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVE 241
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  205 ltspvvaespkkpsmvnnsgkdgntsansadesvkgpnlteiskkitesnavvlavkeve 264

Qy  242 TLLTSIDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAE 301
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  265 tlltsidelakaigkkikndvslidneadhngslisgaylisnlitkkisaikdsgelkae 324

Qy  302 IEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNL 361
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  325 iekakkcseeftaklkgehtdlgkegvtddnakkailktnndktkgadeleklfesvknl 384

Qy  362 SKAAKEMLTNSVKELTS 378
        |||||||||||||||||
Db  385 skaakemltnsvkelts 401
```

RESULT 3
AAB62710
AAB62710 standard; Protein; 374 AA.
AC AAB62710;
XX
DT 03-APR-2001 (first entry)
XX
DE Borrelia sp chimeric ospc protein SEQ ID NO: 24.
KW Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
OS Chimeric - Borrelia sp.
OS Chimeric - Borrelia sp.
XX
PN WO200078966-A1.
XX
PD 28-DEC-2000.
XX
PF 19-JUN-2000; 2000WO-US16915.
XX
PR 18-JUN-1999; 99US-0140042.

XX
PA (UYNY) UNIV NEW YORK STATE RES FOUND.
PA (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI Dattwyler RJ, Selnost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
DR WPI; 2001-050113/06.
DR N-PSDB; AAF29014.
XX
PN WO200078966-A1.

XX
PD 28-DEC-2000.
XX
PF 19-JUN-2000; 2000WO-US16915.
XX
PR 18-JUN-1999; 99US-0140042.
XX
PA (UYNY) UNIV NEW YORK STATE RES FOUND.
PA (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI Dattwyler RJ, Selnost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
DR WPI; 2001-050113/06.
DR N-PSDB; AAF29032.
XX
PT Compositions of OspC polypeptides from strains of Borrelia which cause
PT Lyme disease are used to immunize animals and detect immune responses
PT to Lyme disease -
XX
PS Claim 43; Page 123; 160pp; English.
XX
CC The present invention provides compositions comprising ospC proteins and
CC chimeric ospC proteins from members of the Borrelia genus. These may be
CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC vaccines against Borrelia infection, which is spread by ticks and leads
CC to Lyme disease.
XX
SQ Sequence 398 AA;

Query Match 97.0%; Score 1802.5; DB 22; Length 398;
Best Local Similarity 97.9%; Pred. No. 2.5e-117;
Matches 369; Conservative 3; Mismatches 2; Indels 3; Gaps 1;

Qy 2 ACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKAI 61
       :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2 SCNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKAI 61

Qy 62 GKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNK 121
       |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 62 GKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNK 121

Qy 122 LKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 181
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 122 LKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 181

Qy 182 LTSPVVAESPKKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVE 241
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||
Db 182 LTSPVVAESPKKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVE 241

Qy 242 TLLTSIDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAE 301
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 242 TLLTSIDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAE 301

Qy 302 IEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNL 361
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 302 IEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNL 361

Qy 362 SKAAKEMLTNSVKELTS 378
        |||||||||||||||||
Db 362 SKAAKEMLTNSVKELTS 378

RESULT 5
ID AAB62713 standard; Protein; 377 AA.
XX
AC AAB62713;
XX
DT 03-APR-2001 (first entry)
XX
DE Borrelia sp chimeric ospC protein SEQ ID NO: 30.
XX

RESULT 4
ID AAB62728 standard; Protein; 398 AA.
XX
AC AAB62728;
XX
DT 03-APR-2001 (first entry)
XX
DE Borrelia sp chimeric ospC protein SEQ ID NO: 60.
KW Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS Chimeric - Borrelia sp.
OS Chimeric - Borrelia sp.
XX

KW   Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
XX
PN   WO20078966-A1.
XX
PD   28-DEC-2000.
XX
PF   19-JUN-2000; 2000WO-US16915.
XX
PR   18-JUN-1999;   99US-0140042.
XX
PA   (UYNY) UNIV NEW YORK STATE RES FOUND.
PA   (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI   Dattwyler RJ, Seinost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
PR   WPI: 2001-050113/06.
PR   N-PSDB; AAF29017.
XX

Claim 43; Page 86-87; 160pp; English.

CC   The present invention provides compositions comprising ospc proteins and
CC   chimeric ospc proteins from members of the Borrelia genus. These may be
CC   Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC   vaccines against Borrelia infection, which is spread by ticks and leads
CC   to Lyme disease.
XX
SQ   Sequence   377 AA;

Query Match   86.8%; Score 1614.5; DB 22; Length 377;
Best Local Similarity 87.8%; Pred. No. 2.6e-104;
Matches 331; Conservative 20; Mismatches 25; Indels .1; Gaps 1;

Qy   1 MACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKA 60
         ||||||| |||||||||||||||||||||||||||| | |||:|||||||| ||||||
Db   1 macnnsgkdgntsansadesvkgpnlteinkkitdsnavllavkeveallssideiaaka 60

Qy   61 IGKKIHQNNGLDTENHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN 120
         |||||| |||||||||||||||||||||||||||| | ||||||||||||||||||||
Db   61 igkkihqnngldtenhhgsllagayaistlikqkldglkneglkekidaakkcsetftn 120

Qy 121 KLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 180
         |||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 klkekhtdlgkegvtdadakeailkangtktkgaeelgklfesvevlskaakemlansvk 180

Qy 181 ELTSPVVAESPKKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEV 240
         |||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db 181 eltspvvaespkkpsmvnnsgkdgntsansadesvkgpnlteiskkitesnavvlavkei 240

Qy 241 ETLLTSIDELA-KAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGGELK 299
         ||| ||||||| |||||| |:|||||| | | ||||||||||||:||| ||:||| ||
Db 241 etllasidelatkaigkkiqnngglaveaghngtllagaytiskiitqkidglknsekl 300

Qy 300 AEIEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVK 359
         :|| |||||||||||||||| | || |||| || |||||||:||||||||| ||:||||
Db 301 ekienakkcsedftkklegehaqlgienvtdenakkailtdaakdkgaaelek.lfkave 360

Qy 360 NLSKAAKEMLTNSVKEL 376
         |:|||||||:||||||
Db 361 nlakaakemlansvkel 377

AC   AAB62739;
XX
DT   03-APR-2001 (first entry)
XX
DE   Borrelia sp chimeric ospc protein SEQ ID NO: 82.
XX
KW   Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
XX
PN   WO20078966-A1.
XX
PD   28-DEC-2000.
XX
PF   19-JUN-2000; 2000WO-US16915.
XX
PR   18-JUN-1999;   99US-0140042.
XX
PA   (UYNY) UNIV NEW YORK STATE RES FOUND.
PA   (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI   Dattwyler RJ, Seinost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
PR   WPI: 2001-050113/06.
PR   N-PSDB; AAF29043.
XX

Claim 43; Page 153; 160pp; English.

CC   The present invention provides compositions comprising ospc proteins and
CC   chimeric ospc proteins from members of the Borrelia genus. These may be
CC   Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC   vaccines against Borrelia infection, which is spread by ticks and leads
CC   to Lyme disease.
XX
SQ   Sequence   400 AA;

Query Match   85.5%; Score 1589.5; DB 22; Length 400;
Best Local Similarity 86.7%; Pred. No. 1.5e-102;
Matches 326; Conservative 22; Mismatches 27; Indels 1; Gaps 1;

Qy   2 ACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKAI 61
         :||||||||||||||||||||||||||||||||||||| | |||:||||||||||||||
Db   25 scnnsgkdgntsansadesvkgpnlteinkkitdsnavllavkeveallssideiaakai 84

Qy   62 GKKIHQNNGLDTENHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNK 121
         |||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||
Db   85 gkkihqnngldteynhhgsllagayaistlikqkldglkneglkekidaakkcsetftnk 144

Qy 122 LKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 181
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 145 lkekhtdlgkegvtdadakeailkangtktkgaeelgklfesvevlskaakemlansvke 204

Qy 182 LTSPVVAESPKKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVE 241
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 205 ltspvvaespkkpsmvnnsgkdgntsansadesvkgpnlteiskkitesnavvlavkeie 264

Qy 242 TLLTSIDELA-KAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGGELKA 300
         ||| ||||||| |||||| |||||| | | | ||||||||||||:||| |||||| ||:
Db 265 tllasidelatkaigkkiqnngglaveaghngtllagaytiskiitqkidglknseklke 324

Qy 301 EIEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKN 360
         :|| |||||||||||||||| | || ||||| || |||||||:|||||||||||||||||
Db 325 kienakkcsedftkklegehaqlgienvtdenakkailtdaakdkgaaeleklfkaven 384

Qy 361 LSKAAKEMLTNSVKEL 376
         |:|||||||:||||||
Db 385 lakaakemlansvkel 400

RESULT   6
AAB62739
ID   AAB62739 standard; Protein; 400 AA.
XX

```
RESULT   7
AAB62711
ID   AAB62711 standard; Protein; 373 AA.
XX
AC   AAB62711;
XX
DT   03-APR-2001 (first entry)
XX
DE   Borrelia sp chimeric ospC protein SEQ ID NO: 26.
XX
KW   Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
XX
PN   WO200078966-A1.
XX
PD   28-DEC-2000.
XX
PF   19-JUN-2000; 2000WO-US16915.
XX
PR   18-JUN-1999; 99US-0140042.
XX
PA   (UYNY ) UNIV NEW YORK STATE RES FOUND.
PA   (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI   Dattwyler RJ, Selnost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
DR   WPI; 2001-050113/06.
DR   N-PSDB; AAF29015.
XX
PT   Compositions of OspC polypeptides from strains of Borrelia which cause
PT   Lyme disease are used to immunize animals and detect immune responses
PT   to Lyme disease.
XX
CC   The present invention provides compositions comprising ospC proteins and
CC   chimeric ospC proteins from members of the Borrelia genus. These may be
CC   Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC   vaccines against Borrelia infection, which is spread by ticks and leads
CC   to Lyme disease.
XX
SQ   Sequence    373 AA;

Query Match             84.0%;  Score 1561;  DB 22;  Length 373;
Best Local Similarity   85.9%;  Pred. No. 1.3e-100;
Matches  323;  Conservative   22;  Mismatches   27;  Indels    4;  Gaps    2;

Qy    2 ACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKAI 61
          |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db    1 acnnsgkdgntsansadesvkgpnlteiskkitdsnavllavkeveallssideiaakai 60

Qy   62 GKKIHONNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNK 121
          |||| |||||| ||||||||||||||| || |||||||||||||||||||||||||
Db   61 gkkihqnngldteynhngsllagayaistliqkqkldglkneglkekidaakkcsetftnk 120

Qy  122 LKEKHTDLGKEGVTDADAKERAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 181
          ||||||||||||||||||| |||| ||||||||||||||||||||||||||||||||
Db  121 lkekhtdlgkegvtdadakeailktngtktkgaeelgklfesvevlskaakemlansvke 180

Qy  182 LTSPVVAESPKPKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVE 241
          ||||||||       |||||||||||||||||||||||||||||||||||||||| :1
Db  181 ltspvvaes---panvnsgkdgntsansadesvkgpnlteiskkitesnavvlavkeie 237

Qy  242 TLLTSIDELA-KAIGKKIKNDVSLDNEADHNGSLIGSAYLISNLITKKISAIKDSGELKA 300
          |||| |||| |||||| ||  | | :||||||::  ||: |: :| |:|
Db  238 tilasidelatkaigkkiqgnnglaveaghngtllagaytiskliqkldglknsekike 297

Qy  301 EIEKAKKCSEEFTAKILKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKN 360
```

```
Db  298 klenakkcsedftkklegehaqlgienvtdenakkailicdaakdkqaaeleklfkaven 357
          |||  ||||  ||  ||| || | |   || ||   |||  |||| | ||| ||||  |
Qy  361 LSKAAKEMLTNSVKEL 376
          |: ||||||  |||||
Db  358 lakaakemlansvkel 373


RESULT   8
AAB62729
ID   AAB62729 standard; Protein; 397 AA.
XX
AC   AAB62729;
XX
DT   03-APR-2001 (first entry)
XX
DE   Borrelia sp chimeric ospC protein SEQ ID NO: 62.
XX
KW   Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
XX
PN   WO200078966-A1.
XX
PD   28-DEC-2000.
XX
PF   19-JUN-2000; 2000WO-US16915.
XX
PR   18-JUN-1999; 99US-0140042.
XX
PA   (UYNY ) UNIV NEW YORK STATE RES FOUND.
PA   (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI   Dattwyler RJ, Selnost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
DR   WPI; 2001-050113/06.
DR   N-PSDB; AAF29033.
XX
PT   Compositions of OspC polypeptides from strains of Borrelia which cause
PT   Lyme disease are used to immunize animals and detect immune responses
PT   to Lyme disease.
XX
CC   The present invention provides compositions comprising ospC proteins and
CC   chimeric ospC proteins from members of the Borrelia genus. These may be
CC   Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC   vaccines against Borrelia infection, which is spread by ticks and leads
CC   to Lyme disease.
XX
SQ   Sequence    397 AA;

Query Match             83.8%;  Score 1558;  DB 22;  Length 397;
Best Local Similarity   85.6%;  Pred. No. 2.3e-100;
Matches  322;  Conservative   23;  Mismatches   27;  Indels    4;  Gaps    2;

Qy    2 ACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKAI 61
          ||||||||||||||||||||||||| |||||||||||||||||||||||||||||| ||
Db   25 scnnsgkdgntsansadesvkgpnlteiskkitdsnavllavkeveallssideiaakal 84

Qy   62 GKKIHONNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNK 121
          |||| |||||| ||||||||||||||| || ||||||||||||||||||||||||||
Db   85 gkkihqnngldteynhngsllagayaistliqkqkldglkneglkekidaakkcsetftnk 144

Qy  122 LKEKHTDLGKEGVTDADAKERAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 181
          ||||||||||||||||||| |||| |||||||||||||||||||||||||||||||||
Db  145 lkekhtdlgkegvtdadakeailktngtktkgaeelgklfesvevlskaakemlansvke 204

Qy  182 LTSPVVAESPKPKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVE 241
          ||||||||       ||||||||||||||||||||||||||||||||||||||||| :|
Db  205 ltspvvaes---panvnsgkdgntsansadesvkgpnlteiskkitesnavvlavkeie 261
```

Qy   242 TLLTSIDELA-KAIGKKIKNDVSLDNEADHNGSLISGAYLISNLIITKKISAIKDSGELKA 300
          ||| |||||| ||||||| ::: ::||:::| ||||||| |||:|| ||||||||||||
Db   262 tllasidelatkaigkkiqgnglaveaghngtliagaytiskiiqtkldglknselke   321

Qy   301 EIEKAKKCSEEFTAKILKGEHTDLGKEGVTDDNAKKALIKTNNDKTKGADELEKLFESVKN 360
          :|||||||||:| : ||:|||| ||:|||||:||||| | |: ||| |||||||||:||:|
Db   322 kienakkcsedfkkiegehaqlgienvtdenakkailtdaakdkgaaeleklfkaven   381

Qy   361 LSKAAKEMLTNSVKEL 376
          | |||||| |||||||
Db   382 lakaakemlansvkel 397

RESULT   9
ID   AAB62715 standard; Protein; 370 AA.

DT   03-APR-2001 (first entry)
DE   Borrelia sp chimeric ospc protein SEQ ID NO: 34.
KW   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
PN   WO200078966-A1.
PD   28-DEC-2000.
PF   19-JUN-2000; 2000WO-US16915.
PR   18-JUN-1999; 99US-0140042.
PA   (UYNY) UNIV NEW YORK STATE RES FOUND.
PA   (BROO-) BROOK BIOTECHNOLOGIES INC.
PI   Dattwyler RJ, Seinost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
DR   WPI; 2001-050113/06.
DR   N-PSDB; AAF29019.

PS   Claim 43; Page 91-92; 160pp; English.

PT   Compositions of OspC polypeptides from strains of Borrelia which cause
PT   Lyme disease are used to immunize animals and detect immune responses
PT   to Lyme disease -

The present invention provides compositions comprising ospc proteins and
chimeric ospc proteins from members of the Borrelia genus. These may be
Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
vaccines against Borrelia infection, which is spread by ticks and leads
to Lyme disease.

Sequence 370 AA;

Query Match   82.2%; Score 1528.5; DB 22; Length 370;
Best Local Similarity 86.3%; Pred. No. 2.3e-98;
Matches 327; Conservative 10; Mismatches 31; Indels 11; Gaps   4;

Qy   2 ACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVAKEVEAALLSSIDEIAAKAI   61
          ||||||||||||||||||||||||||||:||||||||||||||||||||||||||| :||||
Db   1 acnnsgkdgntsansadesvkgpnlteiskkitdsnavllavakeveaallssidel-akai   59

Qy   62 GKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLK-NEGLKEKIDAARKCSEIFTN 120
          ||||: :: || : ||| | |||| || | : ||.|| ||||||| |||||| |:|
Db   60 gkkikndgsldneanrneslagaytistilcgklskingseglkekiaaakkcseefst 119

Qy   121 KLKEKHIDLGKEGVTDADAKEALLKANGT-KTKGAEELGKLFESVEVLSKAAKEMLANSV 179

Db   120 klkdnhaqlgigqvtdenakkailkanaagkdkgveelekisgslesl skaakemlansv 179

Qy   180 KELTSPVVAESPKRPSMVNNSGKDGNTSANSADSSVKGPNLTEISKKITESNAVVLAVKE 239
          |||||||| | |||||||||||||||| ||||||||||||||||||||||||||
Db   180 keltspvvhg------nnsgkdgntsansadesvkgpnlteiskkitesnavvlavke   231

Qy   240 VETLLTSIDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELK 299
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   232 vetlltsidelakaigkkikndvsldneadhngslisgaylisnlitkkisaikdsgelk 291

Qy   300 AEIEKAKKCSEEFTAKILKGEHTDLGKEGVTDDNAKKALIKTNNDKTKGADELEKLFESVK 359
          |||||||||||||||| |||||||||||||||||||| |||||||||||||||||||||||
Db   292 aeiekakkcseeftaklkgehtdlgkegvtddnakkailktnndktkgadeleklfesvk   351

Qy   360 NLSKAAKEMLTNSVKELTS 378
          |||||||||||||||||||
Db   352 nlskaakemltnsvkelts 370

RESULT   10
ID   AAB62730 standard; Protein; 394 AA.
AC   AAB62730;

DT   03-APR-2001 (first entry)
DE   Borrelia sp chimeric ospc protein SEQ ID NO: 64.
KW   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
PN   WO200078966-A1.
PD   28-DEC-2000.
PF   19-JUN-2000; 2000WO-US16915.
PR   18-JUN-1999; 99US-0140042.
PA   (UYNY) UNIV NEW YORK STATE RES FOUND.
PA   (BROO-) BROOK BIOTECHNOLOGIES INC.
PI   Dattwyler RJ, Seinost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
DR   WPI; 2001-050113/06.
DR   N-PSDB; AAF29034.

PS   Claim 43; Page 128-129; 160pp; English.

PT   Compositions of OspC polypeptides from strains of Borrelia which cause
PT   Lyme disease are used to immunize animals and detect immune responses
PT   to Lyme disease -

The present invention provides compositions comprising ospc proteins and
chimeric ospc proteins from members of the Borrelia genus. These may be
Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
vaccines against Borrelia infection, which is spread by ticks and leads
to Lyme disease.

Sequence 394 AA;

Query Match   82.1%; Score 1525.5; DB 22; Length 394;
Best Local Similarity 86.0%; Pred. No. 4e-98;
Matches 326; Conservative 11; Mismatches 31; Indels 11; Gaps   4;

Qy   2 ACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVAKEVEAALLSSIDEIAAKAI   61
          :||||||||||||||||||||||||||||:||||||||||||||||||||||||||: |||
Db   25 scnnsgkdgntsansadesvkgpnlteiskkitdsnavllavakeveaallssidel-akai   83

RESULT 11

ID AAB62726 standard; Protein; 384 AA.
XX
AC AAB62726;
XX
DT 03-APR-2001 (first entry)
XX
DE Borrelia sp chimeric ospC protein SEQ ID NO: 56.
XX
KW Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
OS Chimeric - Borrelia sp.
OS Chimeric - Borrelia sp.
XX
PN WO200078966-A1.
XX
PD 28-DEC-2000.
XX
PF 19-JUN-2000; 2000WO-US16915.
XX
PR 18-JUN-1999; 99US-0140042.
XX
PA (UYNY) UNIV NEW YORK STATE RES FOUND.
PA (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI Dattwyler RJ, Seinost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
DR WPI; 2001-050113/06.
DR N-PSDB; AAF29030.
XX
PT Compositions of OspC polypeptides from strains of Borrelia which cause
PT Lyme disease are used to immunize animals and detect immune responses
PT to Lyme disease -
XX
PS Claim 43; Page 117-118; 160pp; English.
XX
CC The present invention provides compositions comprising ospc proteins and
CC chimeric ospC proteins from members of the Borrelia genus. These may be
CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC vaccines against Borrelia infection, which is spread by ticks and leads
CC to Lyme disease.
XX
SQ Sequence 384 AA;

Query Match 80.5%; Score 1496.5; DB 22; Length 384;

```
Qy   62 GKKIHQNNGLDTENNHINGSLLAGAVAYAISTLIKQKLDGLK-NEGLKEKIDAARKKCSEFTN 120
         |||  |||| |||| |||||||||| ||||| |||||  | |||||||||||||||||||| |:
Db   84 gkkikndgsldneanrnesllagaytistltqklskingseglkekiaaakkcseefst 143

Qy  121 KLKEKHTDLGKEGVTDADAKEAILKANGT-KTKGAEELGKLFESVEVLSKAAKEMLANSV 179
         ||| :|   | ||||| ||| |||||| | ||||| ||| | || |||||||||||||||
Db  144 klkdnhaqlqiggvtdenakkallkanaagkdkgveelekisgsleslskaakemlansv 203

Qy  180 KELTSPVVAESPKKPKPSMVVNNGGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKE 239
         ||||||||           |||||||||||||||||||||||||||||||||||||||||||
Db  204 keltspvvhg-------nnsgkdgntsansadesvkgpnlteiskkitesnavvlavke 255

Qy  240 VETLLTTSIDELAKAIGKKIKINDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELK 299
         ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  256 vetlltsidelakaigkkikndvsldneadhngslisgaylisnlitkkisaikdsgelk 315

Qy  300 AEIEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVK 359
         ||||||||||||||||||||||||| |||||||||||| |||||||||||||||:|||||
Db  316 aeiekakkcseeftaklkgehtdlgkegvtddnakkailktnndktkgadelekifesvk 375

Qy  360 NLSKAAKEMLTNSVKELTS 378
         |||||||||||||||||||
Db  376 nlskaakemltnsvkelts 394
```

RESULT 12

ID AAB62737 standard; Protein; 408 AA.
XX
AC AAB62737;
XX
DT 03-APR-2001 (first entry)
XX
DE Borrelia sp chimeric ospC protein SEQ ID NO: 78.
XX
KW Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
OS Chimeric - Borrelia sp.
OS Chimeric - Borrelia sp.
XX
PN WO200078966-A1.
XX
PD 28-DEC-2000.
XX
PF 19-JUN-2000; 2000WO-US16915.
XX
PR 18-JUN-1999; 99US-0140042.
XX
PA (UYNY) UNIV NEW YORK STATE RES FOUND.
PA (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI Dattwyler RJ, Seinost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
DR WPI; 2001-050113/06.
DR N-PSDB; AAF29041.
XX
PT Compositions of OspC polypeptides from strains of Borrelia which cause
PT Lyme disease are used to immunize animals and detect immune responses
PT to Lyme disease -
XX
PS Claim 43; Page 147-148; 160pp; English.
XX
CC The present invention provides compositions comprising ospc proteins and
CC chimeric ospC proteins from members of the Borrelia genus. These may be
CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC vaccines against Borrelia infection, which is spread by ticks and leads

Best Local Similarity 82.0%; Pred. No. 4e-96;
Matches 309; Conservative 32; Mismatches 33; Indels 3; Gaps 2;

```
Qy    2 ACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKAI 61
         |||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db    1 acnnsgkdgntsansadesvkgpnlteiskkitdsnavllavkeveallssideiaakai 60

Qy   62 GKKIHQNNGLDTENNHINGSLLAGAVAYAISTLIKQKLDGLKNEGLKEKIDAARKKCSEFTNK 121
         ||||||| |||| |:|||||||||||| ||||||||||||||||||||||||  |||||||||
Db   61 gkkihqngldteynhngsllagavaistlikqkldglkneglkekiaaakkcseeftnk 120

Qy  122 LKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 181
         |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db  121 lkekhtdlgkegvtdadakeailktgtktkgaeelgklfesvevlskaakemlansvke 180

Qy  182 LTSPVVAESPKKPKPSMVVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKE 241
         ||||||||||| |::|:||||      ||:|  | |:|||||| |:||||:||| :||||||
Db  181 ltspvvaespkpfhgnnsg--gdsastnpdesakgpnltvskkitdsnafilavkeve 238

Qy  242 TLLTSIDELAKAIGKKIKINDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAE 301
         :|| |||||:||||||||:||  | |||| ||||| |||: || |: |||:|| ::  |||
Db  239 allssidelskaigkkikndgtildneanrnesliagayeiskliitqklsvi-nseelkk 297

Qy  302 IEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNL 361
         | :|| |:::|::||| || : || ||  |||||||| || :| |||| || |||:|  |
Db  298 ikeakdcsqkfttkldshaelgigsvqddnakkallkthgtkdkgakeleelfksles 357

Qy  362 SKAAKEMLTNSVKELTS 378
         ||||  :||||||||| :
Db  358 skaaqaaltnsvkeltn 374
```

```
CC   to Lyme disease.
XX
SQ   Sequence  408 AA;
XX

Query Match           80.3%;  Score 1493.5;  DB 22;  Length 408;
Best Local Similarity  81.7%;  Pred. No. 7e-96;
Matches  308;  Conservative  33;  Mismatches  33;  Indels   3;  Gaps   2;

Qy     2 ACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKAI   61
         |:|||||||||||||||||||||||||||||||||||||||||||:|||:||||||||
Db    25 scnnsgkdgntsansadesvkgpnlteiskkitdsnavllavkeveallssideiaakai   84

Qy    62 GKKIHQNNGLDTENNHNGSLLAGAYAISTLLIKQKLDGLKNEGLKEKIDAAKKCSEFFTNK  121
         ||||||||||||||||| |||||||||||||||| ||||||||||||||||||||||
Db    85 gkkihqnngldteynhngsllagayaistlikqkldglkneglkekidaakkcsetfnk   144

Qy   122 LKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE  181
         |||||||||||||||||:|||||||:|||||||||||||||||||||||||||||||
Db   145 lkekhtdlgkegvtdadakeailktngtktkgaeelgklfesveviskaakemlansvke   204

Qy   182 LTSPVVAESPKKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVE  241
         |||||||||||||  ||||  ||  ||||||||||||||||  ||:|||||||||||||
Db   205 ltspvvaespkkpfhgnsg--gdsaastnpdesakgpnltviskkitdsnafllavkeve   262

Qy   242 TLLTSIDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAE  301
         |||:|||||:||:||||||  :||:||||||||||:|:|||  |||||||||||||||
Db   263 allssidelskaigkkikndgtldneanrneslisgayleisklittgktlsvi-nseelkkk   321

Qy   302 IEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNL  361
         |||||:||:::||:||:|||||||||||||||||||||:|:::||:||:|:||:|:|:
Db   322 iekakdcsqkfttkldshaelgigsvqddnakkailkthgtkdkgakelelfkslesl    381

Qy   362 SKAAKEMLTNSVKELTS  378
         ||||||||||||:||:
Db   382 skaaqaaltnsvkeltn  398

RESULT 13
AAB62718
ID   AAB62718 standard; Protein; 368 AA.
XX
AC   AAB62718;
XX
DT   03-APR-2001  (first entry)
XX
```

●

```
DE   Borrelia sp chimeric ospc protein SEQ ID NO: 40.
XX
OS   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
XX
KW   Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
PN   WO200078966-A1.
XX
PD   28-DEC-2000.
XX
PF   19-JUN-2000;  2000WO-US16915.
XX
PR   18-JUN-1999;  99US-0140042.
XX
PA   (UYNY ) UNIV NEW YORK STATE RES FOUND.
PA   (BROO-) BROOK BIOTECHNOLOGIES INC.
PI   Dattwyler RJ, Seinost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
DR   WPI; 2001-050113/06.
DR   N-PSDB; AAF29022.
XX
PT   Compositions of OspC polypeptides from strains of Borrelia which cause
PT   Lyme disease are used to immunize animals and detect immune responses
PT   to Lyme disease -
XX
```

```
PS   Claim 43; Page 99-100; 160pp; English.
XX
CC   The present invention provides compositions comprising ospc proteins and
CC   chimeric ospc proteins from members of the Borrelia genus. These may be
CC   Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC   vaccines against Borrelia infection, which is spread by ticks and leads
CC   to Lyme disease.
XX
SQ   Sequence  368 AA;
XX

Query Match           80.1%;  Score 1489.5;  DB 22;  Length 368;
Best Local Similarity  82.5%;  Pred. No. 1.2e-95;
Matches  311;  Conservative  23;  Mismatches  34;  Indels   9;  Gaps   2;

Qy     2 ACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKAI   61
         |:|||||:||||||||||||||||||||||||||||||||||||||:|||:||||||||
Db     1 acnnsgkdgnasansadesvkgpnlteiskkitesnavllavkeveetllasidelatkai   60

Qy    62 GKKIHQNNGLDTENNHNGSLLAGAYAISTLLIKQKLDGLKNEGLKEKIDAAKKCSEFFTNK  121
         ||||     |:||:||:    |||||||||  ||:|||||:||||||||||||        
Db    61 gkki-gingleanqsknts llsgayaisdliaeklnvlkneelkekidtakqcsteftnk   119

Qy   122 LKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE  181
         ||||:  ||||||||||:|||||:||    ||||||||||:|  ||:||||  ||||||
Db   120 lksehavigldnltddnagrailkhankdkgaaeleklfkavenlskaaqdtlknavke    179

Qy   182 LTSPVVAESPKKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVE  241
         |||:|          ||||||||||||||||||||||||||||||||||||||||||||
Db   180 ltspivhg------nnsgkdgntsansadesvkgpnlteiskkitesnavllavkeve    231

Qy   242 TLLTSIDELAKAIGKKIKNDVSLDNEADINGSLISGAYLISNLITKKISAIKDSGELKAE  301
         |||||||||||||||||||||||||||||||||||||||||||||||:||||||||
Db   232 tlltsidelakaigkkikndvsldneadingslisgayliisnlittkkisaikdsgelkae   301

Qy   302 IEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNL  361
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   302 iekakkcseeftaklkgehtdlgkegvtddnakkailktnndktkgadeleklfesvknl   361

Qy   362 SKAAKEMLTNSVKELTS  378
         |||||||||||||||
Db   362 skaakemltnsvkelts  368

RESULT 14
AAB62734
ID   AAB62734 standard; Protein; 392 AA.
XX
AC   AAB62734;
XX
DT   03-APR-2001  (first entry)
XX
DE   Borrelia sp chimeric ospc protein SEQ ID NO: 72.
XX
OS   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
XX
KW   Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
PN   WO200078966-A1.
XX
PD   28-DEC-2000.
XX
PF   19-JUN-2000;  2000WO-US16915.
XX
PR   18-JUN-1999;  99US-0140042.
XX
PA   (UYNY ) UNIV NEW YORK STATE RES FOUND.
PA   (BROO-) BROOK BIOTECHNOLOGIES INC.
PI   Dattwyler RJ, Seinost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
DR   WPI; 2001-050113/06.
```

DR　N-PSDB; AAF29038.
XX
PT　Compositions of OspC polypeptides from strains of Borrelia which cause
PT　Lyme disease are used to immunize animals and detect immune responses
PT　to Lyme disease -
XX
PS　Claim 43; Page 139-140; 160pp; English.
XX
CC　The present invention provides compositions comprising ospC proteins and
CC　chimeric ospC proteins from members of the Borrelia genus. These may be
CC　Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC　vaccines against Borrelia infection, which is spread by ticks and leads
CC　to Lyme disease.
XX
SQ　Sequence　　392 AA;

Query Match　　80.0%;　Score 1486.5;　DB 22;　Length 392;
Best Local Similarity　82.2%;　Pred. No. 2e-95;
Matches 310;　Conservative　24;　Mismatches　34;　Indels　9;　Gaps　2;

```
Qy       2 ACNNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKAI 61
           :|||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db      25 scnnsgkdgnasansadesvkgpnlteiskkitesnavvlavkeveatllasidelatkai 84

Qy      62 GKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSEFTNK 121
           |||| :||||||||||||||||||||||||||:||:||||||||||||||||||||||||
Db      85 gkki-qnngleangsknts11sgayaisdliaeklnv1kneelkekidtakqcsteftnk 143

Qy     122 LKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 181
           ||||  ||:|| ||||||:|| |||| || ||:|||:| ||||||||||:|| |:|||
Db     144 lksehavlginitddnaqra1ikkhankdkgaaelek1fkavenlskaaqdt1knavke 203

Qy     182 LTSPVVAESPRKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVE 241
           ||||||
Db     182 ltsplvhg-------nnsgkdgntsansadesvkgpn1teiskk1tesnavv1avkeve 255

Qy     204 ltsplvhg------nnsgkdgntsansadesvkgpn1teiskk1tesnavv1avkeve 255

Qy     242 TLLTSTSIDELAKAIGKKIKINDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAE 301
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     256 t11tsidelakaigkk1kndvsldneadhngs1isgaylisnl1tkk1sa1kdsge1kae 315

Qy     302 IEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNL 361
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     316 iekakkcseeftaklkgehtdlgkegvtddnakkailktnndktkgadeleklfesvknl 375

Qy     362 SKAAKEMLTNSVKELTS 378
           |||||||||||||||||
Db     376 skaakemltnsvkelts 392
```

●
RESULT 15
ID　AAB62724
XX
AC　AAB62724;
XX
DT　03-APR-2001　(first entry)
XX
DE　Borrelia sp chimeric ospC protein SEQ ID NO: 52.
XX
KW　Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS　Chimeric - Borrelia sp.
OS　Chimeric - Borrelia sp.
XX
PN　WO200078966-A1.
XX
PD　28-DEC-2000.
XX
PF　19-JUN-2000; 2000WO-US16915.
XX
PR　18-JUN-1999; 99US-0140042.
XX

PA　(UYNY) UNIV NEW YORK STATE RES FOUND.
PA　(BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI　Dattwyler RJ, Seinost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
DR　WPI; 2001-050113/06.
DR　N-PSDB; AAF29028.
XX
PT　Compositions of OspC polypeptides from strains of Borrelia which cause
PT　Lyme disease are used to immunize animals and detect immune responses
PT　to Lyme disease -
XX
PS　Claim 43; Page 112-113; 160pp; English.
XX
CC　The present invention provides compositions comprising ospC proteins and
CC　chimeric ospC proteins from members of the Borrelia genus. These may be
CC　Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC　vaccines against Borrelia infection, which is spread by ticks and leads
CC　to Lyme disease.
XX
SQ　Sequence　　560 AA;

Query Match　　78.6%;　Score 1460.5;　DB 22;　Length 560;
Best Local Similarity　81.7%;　Pred. No. 2e-93;
Matches 309;　Conservative　22;　Mismatches　42;　Indels　5;　Gaps　3;

```
Qy       1 MACNNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKA 60
           ||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||
Db       1 macnnsgkdgntsansadesvkgpn1teiskk1tesnavv1avkeveallssidelaaka 60

Qy      61 IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSEFTTN 120
           |||||:|||||||||||||||||||||||||||||:||:||||||||||||||||||||||
Db      61 igkkihqnngldteynhngsllagayaistlikqkldglkneglkekidaakkcsetftn 120

Qy     121 KLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 klkekhtdlgkegvtdadakeailktngtktkgaeelgklfesvevlskaakemlansvk 180

Qy     181 ELTSPVVAESPKKPSMVNNSGKDGNT-SANSADESVKGPNLTEISKKITESNAVVLAVKE 239
           |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db     181 eltspvvaes---pangsnsgkggdsaastnpadesakgpn1teiskk1tdsnafv1avke 237

Qy     240 VETLLTSIDELA-KAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGEL 297
           |||||:|||||| ||||||||||||||||||||||||||||||||||||:||||:||||
Db     238 vetlvlsidelakaigqkidnnnglaalnnqngsllagayaistl1tekiskiknleel 297

Qy     299 KAEIEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESV 358
           ||| |||||||||||||||||||:|| :|||| |||| |||||||||||||| ||||||
Db     298 kteiakakkcseeftnklksghadigkdatddhakaailkthattdkgakefkdlfesv 357

Qy     359 KNLSKAAKEMLTNSVKEL 376
           : |||||||||||||||
Db     358 eglkaaqvaltnsvkel 375
```

Search completed: March 18, 2002, 09:54:30
Job time: 328 sec

OM protein - protein search, using sw model

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

Run on: March 18, 2002, 09:54:37 ; Search time 124.19 Seconds
(without alignments)
441.679 Million cell updates/sec

Title: US-09-596-746A-24
Perfect score: 1842
Sequence: 1 MACNNSGKDGNTSANSADES.........KNLSKAAKEMLTNSVKELTS 375

Scoring table: BLOSUM62
Searched: 473505 seqs, 146272329 residues
Gapop 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_17:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	926	50.3	192	2	Q9S3P3	borrelia bu
2	921	50.4	192	2	Q9R7B1	borrelia bu
3	910	49.4	191	2	P70818	borrelia bu
4	908	49.3	193	2	Q9RR53	borrelia bu
5	893	48.5	200	2	Q9RQR9	borrelia bu
6	885	48.0	182	2	Q9R7B2	borrelia bu
7	885	48.0	191	2	Q9rqr9	borrelia bu
8	879	47.7	192	2	Q93P2	borrelia bu
9	879	47.7	192	2	031120	borrelia bu
10	853	46.3	177	2	04719	borrelia bu
11	843.5	45.8	178	2	04999	borrelia bu
12	829	45.0	175	2	Q9REH7	borrelia bu
13	802	43.5	163	2	Q9R7B4	borrelia bu
14	786	42.7	165	2	Q9r7a9	borrelia bu
15	773	42.0	159	2	Q9AGB1	borrelia bu
16	757.5	41.1	211	2	Q9R7B0	borrelia bu
17	756.5	41.1	193	2	044720	borrelia bu
18	707	38.4	193	2	Q9237	borrelia bu
19	696.5	37.8	193	2	P94234	borrelia bu
20	683.5	37.1	211	2	049576	borrelia af
21	678.5	36.8	194	2	031122	borrelia bu
22	678.5	36.8	211	2	Q4977	borrelia bu
23	677	36.8	212	2	Q9KIM5	borrelia af
24	676.5	36.7	209	2	Q4671	borrelia af
25	675	36.6	201	2	P96571	borrelia ja
26	674.5	36.6	201	2	044726	borrelia bu
27	674.5	36.6	191	2	P94245	borrelia bu
28	674	36.6	191	2	P96573	borrelia ja
29	673.5	36.6	201	2	044976	borrelia bu
30	672.5	36.5	193	2	P94242	borrelia af
31	672.5	36.5	202	2	P96508	borrelia af
32	672	36.5	203	2	050624	borrelia bu
33	672	36.5	203	2	044705	borrelia bu
34	670	36.4	201	2	P96572	borrelia ja
35	667	36.2	201	2	050619	borrelia bu
36	666	36.2	194	2	P94247	borrelia bu
37	665	36.1	190	2	P94244	borrelia bu
38	661	35.9	212	2	044727	borrelia bu
39	660.5	35.9	201	2	Q9RQR8	borrelia bu
40	658.5	35.7	194	2	031114	borrelia bu
41	658	35.7	193	2	P94231	borrelia bu
42	657.5	35.5	193	2	Q926C7	borrelia bu
43	655.5	35.6	193	2	P94233	borrelia bu
44	655.5	35.6	209	2	044883	borrelia va
45	654.5	35.5	211	2	049577	borrelia af

ALIGNMENTS

RESULT 1
Q9S3P3

```
ID   Q9S3P3      PRELIMINARY;      PRT;   192 AA.
AC   Q9S3P3;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C (FRAGMENT).
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=IP2;
RX   MEDLINE=96296448; PubMed=8709845;
RA   Livey I., Gibbs C.P., Schuster R., Dorner F.;
RT   "Evidence for lateral transfer and recombination in OspC variation in
RT   Lyme disease Borrelia.";
RL   Mol. Microbiol. 18:257-269(1995).
DR   EMBL; L42887; AAB36995.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
FT   NON_TER       1
FT   NON_TER     192
SQ   SEQUENCE   192 AA;  20287 MW;  11846F7AC84C7E3D CRC64;

Query Match             50.3%;  Score 926;  DB 2;  Length 192;
Best Local Similarity   99.5%;
Matches  188;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 CNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAIG 62
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAIG 60

Qy         63 KKIHQNNGLDTEYNHNGSLLAGAYAISTLLIKQKLDGLKNEGLKEKIDAAKKCSEFFTNKL 122
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db         61 KKIHQNNGLDTENNHNGSLLAGAVAISTLLIKQKLDGLKNEGLKEKIDAAKKCSEFFTNKL 120

Qy        123 KEKHTDLGKEGVTDADAKEAILKINGTKIKGAEELGKLFESVEVLSKAAKEMLANSVKEL 182
```

```
Db  121 KEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKEL 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  183 TSPVVAESP 191
            |||||||||
Db  181 TSPVVAESP 189
```

RESULT 2
Q9R7B1
ID   Q9R7B1   PRELIMINARY;   PRT;   192 AA.
AC   Q9R7B1;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C (FRAGMENT).
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TETS;
RX   MEDLINE=97478003; PubMed=9356916;
RA   Ras N.M., Postic D., Foretz M., Baranton G.;
RT   "Borrelia burgdorferi sensu stricto, a bacterial species 'made in the
RT   U.S.A.'?;
RL   Int. J. Syst. Bacteriol. 47:1112-1117(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TETS;
RA   Marti-Ras N., Postic D., Foretz M., Baranton G.;
RL   Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U91798; AAB21895.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
FT   NON_TER   1   1
FT   NON_TER   192   192
SQ   SEQUENCE   192 AA; 20297 MW; 6770502A20AAF764 CRC64;

Query Match   50.0%; Score 921; DB 2; Length 192;
  Best Local Similarity   98.4%;
  Matches 187; Conservative   2; Mismatches   1; Indels   0; Gaps   0;
                               Pred. No. 3, 5e-35;

```
Qy    1 MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA 60
          :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3 ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA 62

Qy   61 IGKKIHONNGLDTEYNHINGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN 120
          ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db   63 IGKKIHONNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN 122

Qy  121 KLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  123 KLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 182

Qy  181 ELTSPVVAES 190
          ||||||||||
Db  183 ELTSPVVAES 192
```

RESULT 3
P70818
ID   P70818   PRELIMINARY;   PRT;   191 AA.
AC   P70818;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN (FRAGMENT).
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=2-1498 CA4;
RA   Probert W.S., Crawford M.R., Cadiz R.B., LeFebvre R.B.;
RL   Submitted (SEP-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; L81131; AAB06569.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
FT   NON_TER   1   1
SQ   SEQUENCE   191 AA; 20126 MW; D2B9B1C82B4DC3C0 CRC64;

Query Match   49.4%; Score 910; DB 2; Length 191;
  Best Local Similarity   98.9%;
  Matches 186; Conservative   0; Mismatches   2; Indels   0; Gaps   0;
                               Pred. No. 1.1e-34;

```
Qy    4 NNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGK 63
          ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 NNSGKGGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGK 60

Qy   64 KIHQNNGLDTEYNHINGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNKLK 123
          |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 KIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNKLK 120

Qy  124 EKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKELT 183
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 EKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKELT 180

Qy  184 SPVVAESP 191
          ||||||||
Db  181 SPVVAESP 188
```

RESULT 4
Q9RR53
ID   Q9RR53   PRELIMINARY;   PRT;   193 AA.
AC   Q9RR53;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C (FRAGMENT).
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=OCI;
RX   MEDLINE=99091544; PubMed=9872945;
RA   Wang I.N., Dykhuizen D.E., Qiu W., Dunn J.J., Bosler E.M., Luft B.J.;
RT   "Genetic diversity of ospC in a local population of Borrelia
RT   burgdorferi sensu stricto.";
RL   Genetics 151:15-30(1999).
DR   EMBL; AF029860; AAB86543.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
FT   NON_TER   1   1
SQ   SEQUENCE   193 AA; 20502 MW; 5EFDB5AFF8986D1E CRC64;

Query Match   49.38%; Score 908; DB 2; Length 193;
  Best Local Similarity   98.4%;
  Matches 184; Conservative   2; Mismatches   1; Indels   0; Gaps   0;
                               Pred. No. 1.3e-34;

```
Qy    1 MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA 60
          :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7 ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA 66

Qy   61 IGKKIHONNGLDTEYNHINGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN 120
```

```
Db       67 IGKKIHQNNGLDTENNHNGSLLAGAYAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSEFFTN 126
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      121 KLKEKHFDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 180

Qy      127 KLKEKHFDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 186

Qy      181 ELTSPVV 187
            |||||||
Db      187 ELTSPVV 193


RESULT 5

Q9RQR9
ID      Q9RQR9            PRELIMINARY;       PRT;    200 AA.
AC      Q9RQR9;
DT      01-MAY-2000 (TrEMBLrel. 13, Created)
DT      01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT      01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE      OUTER SURFACE PROTEIN C PRECURSOR (FRAGMENT).
OSPC.
OC      Borrelia burgdorferi (Lyme disease spirochete).
OC      Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX      NCBI_TaxID=139;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=2E7;
RX      MEDLINE=20002545; PubMed=10531219;
RA      Hofmeister E.K., Glass G.E., Childs J.E., Persing D.H.;
RT      "Population dynamics of a naturally occurring heterogeneous mixture of
RT      Borrelia burgdorferi clones.";
RL      Infect. Immun. 67:5709-5716(1999).
EMBL;   AF074464;   AAD23911.1; -.
DR      InterPro; IPR001800; Lipoprotein_6.
DR      Pfam; PF01441; Lipoprotein_6; 1.
DR      ProDom; PD001149; Lipoprotein_6; 1.
KW      Plasmid.
FT      NON_TER          200
FT      NON_TER          200
SQ      SEQUENCE      200 AA;    21307 MW;    A22F158758BEBB6B CRC64;


Query Match             48.5%;   Score 893;  DB 2;  Length 200;
Best Local Similarity   98.4%;
Matches   181;  Conservative   2;  Mismatches   6.7e-34;   Indels   0;   Gaps   0;
```

(data shown split across the page — sequence alignment block)

```
Qy        1 MACNNSGKDGNTSANSADESVKGPNLTEISKRKITDSNAVLLAVKEVEALLSSIDEIAAKA 60
            :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       17 ISCNNSGKDGNTSANSADESVKGPNLTEISKRKITDSNAVLLAVKEVEALLSSIDEIAAKA 76

Qy       61 IGKKIHQNNGLDTEYNHNGSLLAGAYAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSEFFTN 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       77 IGKKIHQNNGLDTENNHNGSLLAGAYAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSEFFTN 136

Qy      121 KLKEKHFDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 180
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      137 KLKEKHFDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 196

Qy      181 ELTS 184
            ||||
Db      197 ELTS 200
```

```
RESULT 6

Q9R7B2
ID      Q9R7B2            PRELIMINARY;       PRT;    182 AA.
AC      Q9R7B2;
DT      01-MAY-2000 (TrEMBLrel. 13, Created)
DT      01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT      01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE      OUTER SURFACE PROTEIN C (FRAGMENT).
OS      Borrelia burgdorferi (Lyme disease spirochete).
OC      Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX      NCBI_TaxID=139;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=TETS;
RA      Marti-Ras N., Postic D., Foretz M., Baranton G.;
RL      Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.
RN      [2]
RP      SEQUENCE FROM N.A.
RC      STRAIN-TETS;
RA      Marti-Ras N., Postic D., Foretz M., Baranton G.;
RX      MEDLINE=97478003; PubMed=9336916;
RA      Ras N.M., Postic D., Foretz M., Baranton G.;
RT      "Borrelia burgdorferi sensu stricto, a bacterial species 'made in the
RT      U.S.A.'?";
RL      Int. J. Syst. Bacteriol. 47:1112-1117(1997).
OX      NCBI_TaxID=139;
RN      [1]
RP      SEQUENCE FROM N.A.
DR      InterPro; U91797; AAB81894.1; -.
DR      EMBL; U91797; AAB81894.1; -.
DR      Pfam; PF01441; Lipoprotein_6; 1.
DR      ProDom; PD001149; Lipoprotein_6; 1.
FT      NON_TER          182
FT      NON_TER          182
SQ      SEQUENCE      182 AA;    19202 MW;    422146F99AA57BF2 CRC64;


Query Match             48.0%;   Score 885;  DB 2;  Length 182;
Best Local Similarity   99.5%;
Matches   181;  Conservative   0;  Pred. No. 1.4e-33;   Mismatches   1;   Indels   0;   Gaps   0;
```

```
Qy        7 GKDGNTSANSADESVKGPNLTEISKRKITDSNAVLLAVKEVEALLSSIDEIAAKA 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1 GKDGNTSANSADESVKGPNLTEISKRKITDSNAVLLAVKEVEALLSSIDEIAAKAIGKKIH 66

Qy       67 QNNGLDTEYNHNGSLLAGAYAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSEFFTNKLKEKH 126
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1 QNNGLDTENNHNGSLLAGAYAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSEFFTNKLKEKH 60

Qy      127 TDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKELTSPV 186
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61 TDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKELTSPV 120

Qy      187 VA 188
            ||
Db      121 VA 182
```

```
RESULT 7

O31120
ID      O31120            PRELIMINARY;       PRT;    191 AA.
AC      O31120;
DT      01-JAN-1998 (TrEMBLrel. 05, Created)
DT      01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT      01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE      OUTER SURFACE PROTEIN C (FRAGMENT).
OSPC.
OS      Borrelia burgdorferi (Lyme disease spirochete).
OC      Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX      NCBI_TaxID=139;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN-OC10;
RA      Wang I.-N., Dykhuizen D.E., Dunn J.J., Luft B.J.;
RL      Submitted (OCT-1997) to the EMBL/GenBank/DDBJ databases.
DR      EMBL; AF029869; AAB86552.1; -.
DR      InterPro; IPR001800; Lipoprotein_6.
DR      Pfam; PF01441; Lipoprotein_6; 1.
DR      ProDom; PD001149; Lipoprotein_6; 1.
FT      NON_TER          191
FT      NON_TER          191
SQ      SEQUENCE      191 AA;    20502 MW;    E4FB56BC6174DF9 CRC64;


Query Match             48.0%;   Score 885;  DB 2;  Length 191;
Best Local Similarity   100.0%;
Matches   181;  Conservative   0;  Pred. No. 1.5e-33;   Mismatches   0;   Indels   0;   Gaps   0;
```

```
RESULT 8
ID   Q9S3P2       PRELIMINARY;   PRT;   192 AA.
AC   Q9S3P2;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C (FRAGMENT).
GN   OSPC.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=297;
RX   MEDLINE=96296448; PubMed=8709845;
RA   Livey I., Gibbs C.P., Schuster R., Dorner F.;
RT   "Evidence for lateral transfer and recombination in OspC variation in
RT   Lyme disease Borrelia.";
RL   Mol. Microbiol. 18:257-269(1995).
DR   EMBL; L42893; AAB37001.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
NON_TER   1
NON_TER   192
FT
SQ   SEQUENCE   192 AA;   20472 MW;   46AC8F93E4DFED6C CRC64;

Query Match            47.7%;   Score 879;   DB 2;   Length 192;
Best Local Similarity  99.4%;            Pred. No. 2.8e-33;
Matches  180;   Conservative    0;   Mismatches    1;   Indels    0;   Gaps    0;

Qy  195 NNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLTSIDELAKAIGKK 254
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   10 NNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLTSIDELAKAIGKK 69

Qy  255 IKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAEIEKAKKCSEEFTAKLK 314
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   70 IKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAEIEKAKKCSEEFTAKLK 129

Qy  315 GEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLSKAAKEMLTNSVKELT 374
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  130 GEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLSKAAKEMLTNSVKELT 189

Qy  375 S 375
        |
Db  190 S 190
```

```
RESULT 9
ID   Q44719       PRELIMINARY;   PRT;   210 AA.
AC   Q44719;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C (FRAGMENT).
GN   OSPC.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=HB19;
RX   MEDLINE=95154673; PubMed=7851744;
RA   Stevenson B., Barthold S.W.;
RT   "Expression and sequence of outer surface protein C among North
RT   American isolates of Borrelia burgdorferi.";
RL   FEMS Microbiol. Lett. 124:367-372(1994).
DR   EMBL; U04281; AAC43297.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
SQ   SEQUENCE   210 AA;   22526 MW;   2D672991D584E4EE CRC64;

Query Match            47.7%;   Score 879;   DB 2;   Length 210;
Best Local Similarity  99.4%;            Pred. No. 3.1e-33;
Matches  180;   Conservative    0;   Mismatches    1;   Indels    0;   Gaps    0;

Qy  195 NNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLTSIDELAKAIGKK 254
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   20 NNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLTSIDELAKAIGKK 79

Qy  255 IKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAEIEKAKKCSEEFTAKLK 314
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   80 IKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAEIEKAKKCSEEFTAKLK 139

Qy  315 GEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLSKAAKEMLTNSVKELT 374
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  140 GEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLSKAAKEMLTNSVKELT 199

Qy  375 S 375
        |
Db  200 S 200
```

```
RESULT 10
ID   Q44999       PRELIMINARY;   PRT;   177 AA.
AC   Q44999;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C (FRAGMENT).
GN   OSPC.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TXGW;
RX   MEDLINE=95286481; PubMed=7768799;
RA   Theisen M., Borre M., Mathiesen M.J., Mikkelsen B., Lebech A.M.,
RA   Hansen K.;
RT   "Evolution of the Borrelia burgdorferi outer surface protein OspC.";
RL   J. Bacteriol. 177:3036-3044(1995).
DR   EMBL; X84783; CAA59254.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
NON_TER   1
NON_TER   177
FT
SQ   SEQUENCE   177 AA;   18800 MW;   15667189614E7A2D CRC64;

Query Match   46.3%;   Score 853;   DB 2;   Length 177;

Qy  195 NNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLTSIDELAKAIGKK 254
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2 NNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLTSIDELAKAIGKK 61

Qy  255 IKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAEIEKAKKCSEEFTAKLK 314
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   62 IKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAEIEKAKKCSEEFTAKLK 121

Qy  315 GEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLSKAAKEMLTNSVKELT 374
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  122 GEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLSKAAKEMLTNSVKELT 181

Qy  375 S 375
        |
Db  182 S 182
```

RESULT 11

Best Local Similarity 98.9%; Conservative 0; Pred. No. 3.8e-32; Mismatches 2; Indels 0; Gaps 0;
Matches 175; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy    11 NTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGKKIHQNNG 70
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 NTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGKKIHQNNG 60

Qy    71 LDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSEFTFNKLKEKHTDLG 130
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 LDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNKLKEKHTDLG 120

Qy   131 KEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKELTSPVV 187
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 KRGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKELTSPVV 177
```

ID   Q9REH7   PRELIMINARY;   PRT;   178 AA.
AC   Q9REH7;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C (FRAGMENT).
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=B31;
RX   MEDLINE=94075528; PubMed=8253951;
RA   Theisen M., Frederiksen B., Lebech A.M., Vuust J., Hansen K.;
RT   "Polymorphism in ospc gene of Borrelia burgdorferi and
RT   immunoreactivity of OspC protein as a diagnostic antigen.";
RT   use of OspC protein as a diagnostic antigen.";
RL   J. Clin. Microbiol. 31:2570-2576(1993).
RN   [2]
EMBL; X73622; CAA52001.1; -.
InterPro; IPR001800; Lipoprotein_6.
Pfam; PF01441; Lipoprotein_6; 1.
ProDom; PD001149; Lipoprotein_6; 1.
NON_TER 1 1
NON_TER 178 178
SEQUENCE 178 AA; 18894 MW; D619AE4C646F14EB CRC64;

RESULT 12

Query Match 45.8%; Score 843.5; DB 2; Length 178;
Best Local Similarity 98.3%; Pred. No. 1e-31;
Matches 175; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

```
Qy    11 NTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGKKIHQNNG 70
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 NTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAKAIGKKIHQNNG 60

Qy    71 LDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSEFTFNKLKEKHTD-L 129
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 LDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNKLKEKHTDSF 120

Qy   130 GKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKELTSPVV 187
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKELTSPVV 178
```

ID   Q9R7B4   PRELIMINARY;   PRT;   175 AA.
AC   Q9R7B4;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C (FRAGMENT)
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.

RESULT 13

Query Match 45.0%; Score 829; DB 2; Length 175;
Best Local Similarity 98.2%; Pred. No. 4.5e-31;
Matches 167; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```
Qy     1 MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA 60
         :::||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     6 ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA 65

Qy    61 IGKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSEFTFTN 120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    66 IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN 125

Qy   121 KLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKA 170
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   126 KLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKA 175
```

ID   Q9R7A9   PRELIMINARY;   PRT;   163 AA.
AC   Q9R7A9;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C (FRAGMENT)
OS   Borrelia burgdorferi sensu stricto, a bacterial species 'made in the
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TETS;
RX   MEDLINE=97478003; PubMed=9336916;
RA   Marti-Ras N., Postic D., Foretz M., Baranton G.;
RL   Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.
EMBL; U91801; AAB81898.1; -.
InterPro; IPR001800; Lipoprotein_6.
Pfam; PF01441; Lipoprotein_6; 1.
ProDom; PD001149; Lipoprotein_6; 1.
NON_TER 1 1
SEQUENCE 163 AA; 17210 MW; 7550141651BD01FF CRC64;

ID   Q9R7A9   PRELIMINARY;   PRT;   175 AA.
AC   Q9R7A9;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C (FRAGMENT).
OS   Borrelia burgdorferi sensu stricto, a bacterial species 'made in the
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TETS;
RX   MEDLINE=97478003; PubMed=9336916;
RA   Marti-Ras N., Postic D., Foretz M., Baranton G.;
RL   Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.
EMBL; U91792; AAB81889.1; -.
InterPro; IPR001800; Lipoprotein_6.
Pfam; PF01441; Lipoprotein_6; 1.
ProDom; PD001149; Lipoprotein_6; 1.
NON_TER 1 1
SEQUENCE 175 AA; 18573 MW; EC055E7BD3AC3250 CRC64;

RESULT 14
GB1

Query Match
Best Local Similarity   98.8%;   Score 802;   DB 2;   Length 163;
Matches 161; Conservative   1; Mismatches   1; Indels   0; Gaps   0;

```
Qy   2 ACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAI 61
       :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 SCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAI 60

Qy  62 GKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNK 121
       ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db  61 GKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNK 120

Qy 122 LKEKHTDLGKEGVTDADADKEAILKTNGTKTKGAEELGKLFESV 164
       ||||||||||||||||||| ||||||||||||||||||||||||
Db 121 LKEKHTDLGKEGVTDADA KEAILKTNGTKTKGAEELGKLFESV 163
```

ID   Q9AGB1   PRELIMINARY;   PRT;   165 AA.
AC   Q9AGB1;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OSPC-MODIFIED.
OSPC.
GN   OSPC.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=HB19;
RA   Kumaran D., Eswaramoorthy S., Luft B.J., Koide S., Dunn J.J.,
RA   Lawson C.L., Swaminathan S.;
RT   "Crystal structure of outer surface protein C (OspC) from the Lyme
RT   disease spirochete, Borrelia burgdorferi.";
RL   EMBO J. 0:0-0(2001).
DR   EMBL; AF337548; AAK21289.1; -.
FT   VARIANT   61   61   M -> I.
SQ   SEQUENCE   165 AA;   17833 MW;   56E01536D22F61BF CRC64;

RESULT 15
ID   Q9R7B0   PRELIMINARY;   PRT;   159 AA.
AC   Q9R7B0;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C (FRAGMENT).
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TETS;

Query Match
Best Local Similarity   98.2%;   Score 786;   DB 2;   Length 165;
Matches 161; Conservative   2; Mismatches   1; Indels   0; Gaps   0;

```
Qy 212 VKGPNLTEISKKITESNAVLAVKEVETLLTSIDELAKAIGKKIKNDVSLDNEADHNGSL 271
       :||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||
Db   1 MKGPNLTEISKKITESNAVVLAVKEVETLLTSIDELAKAIGKKIKNDVSLDNEADHNGSL 60

Qy 272 ISGAYLISNLITKKISAIKDSGELKAEIEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAK 331
       :|||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||
Db  61 MSGAYLISTLITKKISAIKDSGELKAEIEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAK 120

Qy 332 KAILKTNNDKTKGADELEKLFESVKNLSKAAKEMLTNSVKELTS 375
       ||||||||||||||||||||||||||||||||||||||||||||
Db 121 KAILKTNNDKTKGADELEKLFESVKNLSKAAKEMLTNSVKELTS 164
```

RESULT 13
GB1

Query Match
Best Local Similarity   99.4%;   Score 773;   DB 2;   Length 159;
Matches 158; Conservative   0; Mismatches   1; Indels   0; Gaps   0;

```
Qy  12 TSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGKKIHQNNGL 71
       |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 TSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGKKIHQNNGL 60

Qy  72 DTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNKLKEKHTDLGK 131
       ||||.||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  61 DTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNKLKEKHTDLGK 120

Qy 132 EGVTDADADKEAILKTNGTKTKGAEELGKLFESVEVLSKA 170
       |||||||| |||||||||||||||||||||||||||||||
Db 121 EGVTDADA KEAILKTNGTKTKGAEELGKLFESVEVLSKA 159
```

RX   MEDLINE=97478003; PubMed=9336916;
RA   Ras N.M., Postic D., Foretz M., Baranton G.;
RT   "Borrelia burgdorferi sensu stricto, a bacterial species 'made in the
U.S.A.'?";
RL   Int. J. Syst. Bacteriol. 47:1112-1117(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TETS;
RA   Marti-Ras N., Postic D., Foretz M., Baranton G.;
RL   Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U91799; AAB18946.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
NON_TER   1   1
NON_TER   159   159
SQ   SEQUENCE   159 AA;   16861 MW;   5285D422C43DC213 CRC64;

THIS PAGE BLANK (USPTO)

```
                                          GenCore version 4.5
                               Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on:            March 18, 2002, 09:55:42 ; Search time 39.62 Seconds
                                              (without alignments)
                                              347.030 Million cell updates/sec

Title:             US-09-596-746A-24.
Perfect score:     1842
Sequence:          1 MACNNSGKDGNTSANSADES............KNLSKAAKEMLTNSVKELTS 375

Scoring table:     BLOSUM62
                   Gapop 10.0 , Gapext 0.5

Searched:          100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters:          100059

Post-processing:   Minimum Match 0%
                   Maximum Match 100%
                   Listing first 45 summaries

Database :         SwissProt_39:*

                   Pred. No. is the number of results predicted by chance to have a
                   score greater than or equal to the score of the result being printed,
                   and is derived by analysis of the total score distribution.

                                         SUMMARIES

Result                   %
No.   Score   Query Match Length DB  ID            Description
----  -----   ----------------------  --            -----------
1     928     50.4   210   1  OSC1_BORBU    borrelia bu
2     614     33.3   212   1  OSC2_BORBU    borrelia bu
3     404.5   22.0   212   1  VM24_BORHE    borrelia he
4     214     21.7   214   1  VM03_BORHE    borrelia he
5     176.5   9.6    1509  1  MYSN_ACACA    acanthamoeb
6     170     9.2    1251  1  RBP2_PLAVB    plasmodium
7     169     9.2    1957  1  YD86_SCHPO    schizosacch
8     167     9.1    483   1  M6_STRPY      streptococc
9     167     9.1    564   1  M12_STRPY     streptococc
10    163.5   8.9    1790  1  USO1_YEAST    saccharomyc
11    162.5   8.8    3210  1  CENF_HUMAN    homo sapien
12    159     8.6    2116  1  MYS2_DICDI    dictyosteli
13    157.5   8.5    1786  1  LMB1_MOUSE    mus musculu
14    157     8.5    998   1  SCA4_RICAK    rickettsia
15    157     8.5    1938  1  MYS_AEQIR     aequipecten
16    155     8.4    705   1  YNP9_CAEEL    caenorhabdi
17    154     8.4    1966  1  MYSB_CAEEL    caenorhabdi
18    153     8.3    1427  1  REST_HUMAN    homo sapien
19    152     8.3    492   1  M5_STRPY      streptococc
20    150.5   8.2    775   1  YHGE_BACSU    bacillus su
21    148.5   8.1    1713  1  LMA3_HUMAN    homo sapien
22    148     8.0    1433  1  REST_CHICK    homo sapien
23    147     8.0    1164  1  BAG_STRAG     galius gall
24    146.5   8.0    875   1  ZIP1_YEAST    saccharomyc
25    146     7.9    1786  1  LMB1_HUMAN    homo sapien
26    146     7.9    2349  1  TPR_HUMAN     homo sapien
27    145.5   7.9    1500  1  SSP5_STRGN    streptococc
28    145.5   7.9    1961  1  MYH9_RAT      rattus norv
29    145     7.9    444   1  PUR2_METJA    methanococc
30    145     7.9    962   1  VDP_HUMAN     homo sapien
31    145     7.9    2869  1  RBP1_PLAVB    plasmodium
32    145     7.9    3110  1  LMA2_HUMAN    homo sapien
33    144     7.8    1189  1  SCI1_CHICK    gallus gall
34    144     7.8    2017  1  MYSN_DROME    drosophila
35    142.5   7.7    679   1  YKR9_YEAST    saccharomyc
36    142.5   7.7    3106  1  LMA2_MOUSE    mus musculu
37    142     7.7    1084  1  MYSS_RABIT    oryctolagus
38    142     7.7    1935  1  MYSS_CYPCA    cyprinus ca
39    141.5   7.7    961   1  VDP_BOVIN     bos taurus
40    141.5   7.7    1013  1  SCA4_RICRH    rickettsia
41    141     7.7    1186  1  O9aJ81        ricketsia
42    140.5   7.7    1067  1  SMC_BACSU     bacillus su
43    140     7.6    1312  1  EG52_XENLA    xenopus lae
44    140     7.6    1060  1  RA0_YEAST     saccharomyc
45    139.5   7.6    1526  1  EG51_XENLA    xenopus lae
                           P28025  saccharomyc
                           P12753  xenopus lae
                           Q9usi6  schizosacch
```

                                         ALIGNMENTS

```
RESULT
OSC1_BORBU

1
ID   OSC1_BORBU       STANDARD;      PRT;      210 AA.
AC   Q07337;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   OUTER SURFACE PROTEIN C PRECURSOR (PC).
GN   OSPC OR BBB19.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OG   Plasmid lp54.
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 35210 / B31;
RX   MEDLINE=93268136; PubMed=8098841;
RA   Jauris-Heipke S., Fuchs R., Motz M., Preac-Mursic V., Schwab E.,
RA   Will G., Wilske B.;
RT   "Genetic heterogenity of the genes coding for the outer surface
RT   protein C (OspC) and the flagellin of Borrelia burgdorferi.";
RL   Med. Microbiol. Immunol. 182:37-50(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 35210 / B31;
RX   MEDLINE=93239332; PubMed=8478108;
RA   Wilske B., Preac-Mursic V., Jauris S., Pradel I., Soutschek E.,
RA   Schwab E., Wanner G.;
RT   "Immunological and molecular polymorphisms of OspC, an immunodominant
RT   major outer surface protein of Borrelia burgdorferi.";
RL   Infect. Immun. 61:2182-2191(1993).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 35210 / B31;
RX   MEDLINE=96025162; PubMed=7494039;
RA   Fukunaga M., Hamase A.;
RT   "Outer surface protein C gene sequence analysis of Borrelia
RT   burgdorferi sensu lato isolates from Japan.";
RL   J. Clin. Microbiol. 33:2415-2420(1995).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 35210 / B31;
RX   MEDLINE=94041630; PubMed=8225587;
RA   Padula S.J., Sampieri A., Dias F., Sczepanski A., Ryan R.W.;
RT   "Molecular characterization and expression of p23 (OspC) from a North
RT   American strain of Borrelia burgdorferi.";
RL   Infect. Immun. 61:5097-5105(1993).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 35210 / B31;
RX   MEDLINE=98065943; PubMed=9403685;
RA   Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA   Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
RA   Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
RA   Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,
RA   van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
```

RA   Utterback T., Watthey L., McDonald L., Artiach P., Bowman C.,
RA   Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
RA   Smith H.O., Venter J.C.;
RT   "Genomic sequence of a Lyme disease spirochaete, Borrelia
RT   burgdorferi.";
RL   Nature 390:580-586(1997).
CC   -!- FUNCTION: NOT KNOWN; MAJOR IMMUNODOMINANT PROTEIN.
CC   -!- SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID
CC   ANCHOR.
CC
DR   EMBL; X69596; CAA49306.1; -.
DR   EMBL; U01894; AAA16058.1; -.
DR   EMBL; D49497; BAA08457.1; -.
DR   EMBL; AE000792; AAC66329.1; -.
DR   TIGR; BBB19; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
DR   PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KW   Outer membrane; Lipoprotein; Signal; Plasmid; Antigen;
KW   Complete proteome.
FT   SIGNAL   1   18   BY SIMILARITY.
FT   CHAIN   19   210   OUTER SURFACE PROTEIN C.
FT   LIPID   19   19   N-ACYL DIGLYCERIDE (BY SIMILARITY).
SQ   SEQUENCE   210 AA; 22340 MW; 7A4FC978F91777BF CRC64;

Query Match   50.4%; Score 928; DB 1; Length 210;
Best Local Similarity   98.4%;
Matches 188; Conservative   2; Mismatches   1; Indels   0; Gaps   0;

Qy   1 MACNNSGKDGNTSANSADESVKGPNLTETSKKITDSNAVLLAVKEVEALLSSIDEIARAKA 60
          :|||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db   17 ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIARAKA 76

Qy   61 IGKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSEIFTN 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   77 IGKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSEIFTN 136

Qy   121 KLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   137 KLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 196

Qy   181 ELTSPVVAESP 191
          |||||||||||
Db   197 ELTSPVVAESP 207

RESULT 2
ID   OSC2_BORBU   STANDARD;   PRT;   212 AA.
AC   Q08137;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   OUTER SURFACE PROTEIN C PRECURSOR (PC).
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
RX   MEDLINE=92219995; PubMed=1560779;
RA   Fuchs R., Jauris S., Lottspeich F., Preac-Mursic V., Wilske B.,
RA   Soutschek E.;
RT   "Molecular analysis and expression of a Borrelia burgdorferi gene
RT   encoding a 22 kDa protein (pC) in Escherichia coli.";
RL   Mol. Microbiol. 6:503-509(1992).
RN   [2]
RP   SEQUENCE OF 1-205 FROM N.A.
RX   STRAIN=DK26;
RX   MEDLINE=94075528; PubMed=8253951;
RA   Theisen M., Frederiksen B., Lebech A.M., Vuust J., Hansen K.;
RT   "Polymorphism in OspC gene of Borrelia burgdorferi and
RT   immunoreactivity of OspC protein: Implications for taxonomy and for
RT   use of OspC protein as a diagnostic antigen.";
RL   J. Clin. Microbiol. 31:2570-2576(1993).
CC   -!- FUNCTION: NOT KNOWN; MAJOR IMMUNODOMINANT PROTEIN.
CC   -!- SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID
CC   ANCHOR.
CC
DR   EMBL; X62162; CAA44093.1; -.
DR   EMBL; X73624; CAA52003.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
KW   Outer membrane; Lipoprotein; Signal; Plasmid; Antigen.
FT   SIGNAL   1   18   BY SIMILARITY.
FT   CHAIN   19   212   OUTER SURFACE PROTEIN C.
FT   LIPID   19   19   N-ACYL DIGLYCERIDE (BY SIMILARITY).
SQ   SEQUENCE   212 AA; 22499 MW; C206C231FBF2E7D4 CRC64;

Query Match   33.3%; Score 614; DB 1; Length 212;
Best Local Similarity   69.9%; Conservative
Matches 135; Pred. No. 1.4e-21; Mismatches 42; Indels 2; Gaps 2;

Qy   1 MACNNSGKDGNT-SANSADESVKGPNLTETSKKITDSNAVLLAVKEVEALLSSIDEIARAK 59
          :||||||||:|| ||||||::|||||||:||||||||||:||||||||||||||:|:|||
Db   17 ISCNNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVEALLSSIDELARK 76

Qy   60 AIGKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKN-EGLKEKIDAAKKCSETF 118
          ||:||| :||||| :||:|||||||||||||||: :||: |||| ::| :|:||:|||| |
Db   77 AIGQKIDNNNGLAALNNQGSLLAGAYAISTLITEKLSKLKNLEELKTEIAKAKKCSEEF 136

Qy   119 TNKLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANS 178
          |||||.   :|||: || ||:|||| ::| ||| :|||||||||| ||||:: ||:||
Db   137 TNKLKSGHADLGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGLLKAAQVALTNS 196

Qy   179 VKELTSPVVAESP 191
          ||||||||||||
Db   197 VKELTSPVVAESP 209

RESULT 3
ID   VM24_BORHE   STANDARD;   PRT;   214 AA.
AC   P32778;
DT   01-OCT-1993 (Rel. 27, Created)
DT   01-OCT-1993 (Rel. 27, Last sequence update)
DT   01-FEB-1994 (Rel. 28, Last annotation update)
DE   VARIABLE MAJOR OUTER MEMBRANE LIPOPROTEIN 24 PRECURSOR.
DE   VMP24.
OS   Borrelia hermsii.
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=140;
RN   [1]
RP   SEQUENCE FROM N.A.

RC   STRAIN=SSP, HS1 SEROTYPE 24;
RX   MEDLINE=93133110; PubMed=1484486;
RA   Restrepo B.I., Kitten T., Carter C.J., Infante D., Barbour A.G.;
RT   "Subtelomeric expression regions of Borrelia hermsii linear plasmids
RT   are highly polymorphic.";
RL   Mol. Microbiol. 6:3299-3311(1992).
CC   -!- FUNCTION: SERVES TO AVOID THE HOST IMMUNE RESPONSE BY CHANGING
CC   FROM ONE SURFACE EXPOSED VMP TO ANOTHER.
CC   -!- SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID
CC   ANCHOR.
CC   -!- SIMILARITY: STRONG, TO VMP3.
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --
DR   EMBL; L04786; AAA22964.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
DR   PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KW   Outer membrane; Lipoprotein; Signal; Plasmid.
FT   SIGNAL   1   18   PROBABLE.
FT   CHAIN   19   214   VARIABLE MAJOR OUTER MEMBRANE
FT   LIPOPROTEIN 24.
FT   LIPID   19   19   N-ACYL DIGLYCERIDE (PROBABLE).
SQ   SEQUENCE   214 AA; 22541 MW; F1583F510246F7C7 CRC64;

Query Match   22.0%; Score 404.5; DB 1; Length 214;
Best Local Similarity   45.4%;
Matches   93; Conservative   36; Mismatches   59; Indels   17; Gaps   5;

Qy   1 MACNNSGKDGNTSANSADESVKGP-----NLTEISKKITDSNAVLLAVKEVEALLSSIDEI 56
          ||| || |:|| || ||| |: || |:||:||| ::|||||| :|||||||||:|
Db   17 MSCNNGGPE-----LKSDEVAKSDGTVLDLAKVSKKKIKEASAFAASVKEVETLVKSVDEL 71

Qy   57 AAKAIGKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKOKL------DGLKNEGLKEKIDA 110
          ||||| ||| | || | ||| ||| ::||| || | :| |||| |:||
Db   72 -AKAIGKKIKNDGGLDTEAGONGSLIAGVHSVVSAVKIKVGALETTSGISNE-LKTKITE 129

Qy   111 AKKCSEIFTNKLKEKHIDLGKESGVTDADAKEAILKTNGTKTKTGAEELGKLFESVEVLSKA 170
          |||| | |||| | |:| ::||| | |:| |||||| | |||||||:||||||||||
Db   130 VKSKAEAFLNKLKDGHTELGKDDASDDTTKKAIKKDNSDKTKGASELEALNTAVDALLKA 189

Qy   171 AKEMLANSVKELTSPVVAESPAMVN 195
          ||:||||||||||||:||:| | |
Db   190 AEGEVEAAIKELTAPVKAEKPSQNN 214

RT   are highly polymorphic.";
RL   Mol. Microbiol. 6:3299-3311(1992).
CC   -!- FUNCTION: SERVES TO AVOID THE HOST IMMUNE RESPONSE BY CHANGING
CC   FROM ONE SURFACE EXPOSED VMP TO ANOTHER.
CC   -!- SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID
CC   ANCHOR.
CC   -!- SIMILARITY: STRONG, TO VMP24.
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --
DR   EMBL; L04789; AAA22967.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
DR   PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KW   Outer membrane; Lipoprotein; Signal; Plasmid.
FT   SIGNAL   1   18   PROBABLE.
FT   CHAIN   19   215   VARIABLE MAJOR OUTER MEMBRANE
FT   LIPOPROTEIN 3.
FT   LIPID   19   19   N-ACYL DIGLYCERIDE (PROBABLE).
SQ   SEQUENCE   215 AA; 23139 MW; 684C74D35F87C771 CRC64;

Query Match   21.7%; Score 400; DB 1; Length 215;
Best Local Similarity   45.9%;
Matches   94; Conservative   34; Mismatches   61; Indels   16; Gaps   5;

Qy   1 MACNNSGKDGNTSANSADESVKGP-----NLTEISKKITDSNAVLLAVKEVEALLSSIDEI 56
          ||| || |:|| || ||| |: || |:||:||| ::|||||| :|||||||||:|
Db   17 MSCNNGGPE-----LKSDEVAKSDGTVLDLAKVSKKKIKDASDFAASVKEVHTLVKSIDEL 71

Qy   57 AAKAIGKKI-HQNNGLDTEYNHNGSLLAGAYAISTLIKQKL------DGLK------NEGLKEKIDA 110
          ||||| ||| :| || | ||| ||| ::|| | | :| :| | |:||
Db   72 -AKAIGKKIKNDNSNEFEDENDHNGSLIAGVFQVILTVKAKLTSLEQIIGISDELKTEVGM 130

Qy   111 AKKCSEIFTNKLKEKHIDLGKEGVTDADAKEAILKTNGTKTKTGAEELGKLFESVEVLSKA 170
          |||| | |||| | |:| ::||| | |:| |||||| | |||||||:||||||||||
Db   131 VKKESEAFVTQVKSKHHTDLAKEGVTDAHAKSAILVTDGTKDKGAAAELIKLNTAIDELLKA 190

Qy   171 AKEMLANSVKELTSPVVAESPAMVN 195
          ||:||||||||||||:||:| | |
Db   191 ANDAVETVIKELTASVKAEKPSQNN 215

RESULT   4
ID   VM03_BORHE   STANDARD;   PRT;   215 AA.
AC   O02448;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   01-FEB-1994 (Rel. 28, Last annotation update)
DE   VARIABLE MAJOR OUTER MEMBRANE LIPOPROTEIN 3 PRECURSOR.
GN   VMP3.
OS   Borrelia hermsii.
OG   Plasmid.
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=140;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SSP, HS1 SEROTYPE 3;
RX   MEDLINE=93133110; PubMed=1484486;
RA   Restrepo B.I., Kitten T., Carter C.J., Infante D., Barbour A.G.;
RT   "Subtelomeric expression regions of Borrelia hermsii linear plasmids

---

RESULT   5
ID   MYSN_ACACA   STANDARD;   PRT;   1509 AA.
AC   P05659;
DT   01-NOV-1988 (Rel. 09, Created)
DT   01-NOV-1988 (Rel. 09, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   MYOSIN II HEAVY CHAIN, NON MUSCLE.
OS   Acanthamoeba castellanii (Amoeba).
OC   Eukaryota; Acanthamoebidae; Acanthamoeba.
OX   NCBI_TaxID=5755;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87308395; PubMed=3040773;
RA   Hammer J.A. III, Bowers B., Paterson B.M., Korn E.D.;
RT   "Complete nucleotide sequence and deduced polypeptide sequence of a
RT   nonmuscle myosin heavy chain gene from Acanthamoeba: evidence of a
RT   hinge in the rodlike tail.";
RL   J. Cell Biol. 105:913-925(1987).
CC   -!- FUNCTION: MYOSIN IS A PROTEIN THAT BINDS TO F-ACTIN & HAS ATPASE
CC   ACTIVITY THAT IS ACTIVATED BY F-ACTIN.
CC   -!- SUBUNIT: MYOSIN II HEAVY CHAIN IS TWO-HEADED. IT SELF-ASSEMBLES
CC   INTO FILAMENTS. HEXAMER OF 2 HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI

Left column

```
        LIGHT CHAIN SUBUNITS (MLC) AND 2 REGULATORY LIGHT CHAIN SUBUNITS
        (MLC-2).
-!- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
    CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
    CHARACTERISTIC OF ALPHA-HELICAL STRUCTURES. THIS REGION IS
    INTERRUPTED BY A HINGE AND JOINED BY A NONHELICAL TAILPIECE WHERE
    THE REGULATORY PHOSPHORYLATION SITES RESIDE.
-!- MISCELLANEOUS: THE HINGE REGION MAY PLAY A KEY ROLE IN MEDIATING
    THE EFFECT OF HEAVY CHAIN PHOSPHORYLATION ON ENZYMATIC ACTIVITY.
-!- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
-!- SIMILARITY: CONTAINS 2 IQ DOMAINS.
```

```
DR   EMBL; Y00624; CAA68663.1; -.
DR   PIR; A27224; A27224.
DR   HSSP; P08799; 1MND.
DR   InterPro; IPR000048; IQ.
DR   InterPro; IPR001609; myosin_head.
DR   Pfam; PF00612; IQ; 2.
DR   Pfam; PF00063; myosin_head; 1.
DR   PRINTS; PR00193; MYOSINHEAVY.
DR   ProDom; PD000355; myosin_head; 1.
DR   SMART; SM0015; IQ; 1.
DR   SMART; SM00242; MYSc; 1.
DR   PROSITE; PS50096; IQ; 1.
KW   Myosin; Coiled coil; Actin-binding; ATP-binding; Calmodulin-binding;
KW   Methylation; Alkylation; Phosphorylation; Multigene family.
FT   DOMAIN        1    789       MYOSIN HEAD-LIKE.
FT   DOMAIN      790    819       IQ.
FT   DOMAIN      848   1509       COILED COIL (POTENTIAL).
FT   DOMAIN      848   1226       ALPHA-HELICAL TAILPIECE (S2).
FT   DOMAIN     1227   1252       HINGE.
FT   DOMAIN     1253   1509       LIGHT MEROMYOSIN (LMM).
FT   DOMAIN     1253   1482       ALPHA-HELICAL TAILPIECE (LMM).
FT   DOMAIN     1483   1509       NONHELICAL TAILPIECE (LMM).
FT   NP_BIND     182    189       ATP.
FT   DOMAIN      660    682       ACTIN-BINDING.
FT   DOMAIN      766    789       ACTIN-BINDING.
FT   MOD_RES     133    133       METHYLATION (TRI-) (POTENTIAL).
FT   MOD_RES     700    700       ALKYLATION (SH-1) (POTENTIAL).
FT   MOD_RES    1489   1489       PHOSPHORYLATION.
FT   MOD_RES    1494   1494       PHOSPHORYLATION.
FT   MOD_RES    1499   1499       PHOSPHORYLATION.
SQ   SEQUENCE   1509 AA;  171201 MW;  2CE49BE51173D17E CRC64;

Query Match
Best Local Similarity   9.68%;
Matches  103;  Conservative  68;

Score 176.5;  DB 1;  Length 1509;
Pred. No. 0.41;  Mismatches 137;  Indels 107;  Gaps 21;

Qy   226 -----ESNAVVLAVKEVETLLTSIDEL-------AKAIGKKIKNDVSLDNEADHNGSL 271
              :   |      | :: ||| |  :: |   :| :  || :| |  || :::||
Db  1161 ALDGQKNA---AAAQOAKTLKTQVDETKRRLEEAEASAARLEKERKN--ALDEVAQ---- 1210

Qy   272 ISGAYLISNLITKISAIKDSG------------ELKAETEKAKK---CSEEFTAKLRGE 316
              :|     | :: |  :||         |:  ::|| |  || |  ::||
Db  1211 --------LTADLDAERDSGAQORRKLNTRISELQSELENAPKTGGASSEEVKRLEGE 1260

Qy   317 HTDLGKEGVTDDNAKKAILKTNDDKTKGADELERKLFESVYKNLSKAAKEMLTNSVK 371
         ||:  |   ||:| |||| ||:| ||| ||| :|||||| :| ||||||| |: ::|
Db  1261 LERLEEELLTAQEARAAAEK-NLDKAN--LELEELRQEADDAARDNDKLVKDNRK 1312

Qy    28 EISKKITDSNAVLLAVKEVEA------LLSSIDE-----IAAKAIGKKIHONNGLDTEYNH 77
            |:|  ::::  | | || :::|  |  || :::| : :|
Db   934 ELQEETSASNDILEQKRKLEAEKGELKASLEEEEENRKALQEAKTKVESERNELQDKYED 993

Qy    78 NGSLLAGAYAISTLLIKQKLDGLKNEGLKEKIDA------AKKCSETFTNKLIKEKHTDLGKEGV 134
         |:|||| :  ::| || :||:    |: :||| || ::   |  :  || :|||| ||:  |
Db   994 EA-----AAHDSLKKEEDLSRE-LRETKDALADAENISETLRSKL--KNTERGADDV 1043

Qy   135 TDA--DAKEAILKTNGTKTRGAEELGKLFESVEVLSKAAKEMLANSVKE------ 181
          :|  || | | | ||:|  |  ||| :|| :|| :| ||||||| :|
Db  1044 RNELDDVTATKLQLEKTKKSLEEELAQTRAQLEE-EKSGKEAASSAKAKQLGQOLEDARSE 1102

Qy   182 ---LTSPVVAESPAMVNNSGKDGN------------TSANSADESVKG--PNLTEISKKIT 225
             | |    |||   | ::|  |        |:| | | :: |  | :
Db  1103 VDSLKSKLSAAEKSL--KTAKDQNRDLDEQLEDERTVRANVDKQKKALEAKLTELEQOVT 1160
```

Right column

```
RESULT   6
RBP2_PLAVB

ID   RBP2_PLAVB     STANDARD;      PRT;   1251 AA.
AC   Q00799;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-APR-1993 (Rel. 25, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   RETICULOCYTE BINDING PROTEIN 2 (FRAGMENT).
GN   RBP2.
OS   Plasmodium vivax (strain Belem).
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=31273;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92315338; PubMed=1617731;
RA   Galinski M.R., Medina C.C., Ingravallo P., Barnwell J.W.;
RT   "A reticulocyte-binding protein complex of Plasmodium vivax
RT   merozoites.";
RL   Cell 69:1213-1226(1992).
CC   -!- FUNCTION: INVOLVED IN RETICULOCYTE ADHESION. SPECIFICALLY BINDS TO
CC       HUMAN RETICULOCYTE CELLS.
CC   -!- SUBCELLULAR LOCATION: MEMBRANE-BOUND (PROBABLE).
```

```
DR   EMBL; M88098; AAA29744.1; -.
KW   Malaria; Receptor; Membrane.
FT   NON_TER        1      1
FT   NON_TER     1251   1251
SQ   SEQUENCE   1251 AA;  143741 MW;  54BA51C7404AC572 CRC64;

Query Match
Best Local Similarity  22.9%;
Matches  93;  Conservative  50;
Score 170;  DB 1;  Length 1251;
Pred. No. 0.65;  Mismatches 139;  Indels 124;  Gaps 17;

Qy    22 KGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA--IGKKIHONNGLDTEYNHNG 79
          |   |: | |:|| ::   | :  :|  :| :        || || :
Db   774 KNTNELDVHKNIQDAYKVALEI-----LAHSDEIDTKQKDSSKLIEMGNQI------- 819

Qy    80 SILAGAYAISTLLIKQKLDGLKNEGLKEKIDAAKKCSETFTNKL---KEKHTDLGKEGVTD 136
          ::| | |:||| ||:|                |:|  || :||     |:||:  |: :
Db   820 -----YLKVVLINQ------YKNKISSIKSKEEAVSVKIGNVSKKHSELSKITCSD 864

Qy   137 ADAKEAIL-------KINGTKTKGAEELGKLFESVEVLSKAAKEMLANSV 179
          | |  |         |:| |:|||||| :||| :|| :| ||||||||:
Db   865 KSYDNIIALEKQTELQNLRNSFTQEKINTNSDSKLEKIKTDFES------ 913

Qy   180 KELTSPVVAESPAMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVET 239
          | | |       | ::|  |         :|| ||:|| | |:||:| ||:: | |: :
Db   914 KILEGEVNA-----LKASSDNHEHVQSKSEPV-NPALSEIEKEET------- 956

Qy   240 LLTSIDELAK--------------AIGKKIKNDVSLDNEADHNGSLISGAYL------- 277
          |||:|:|||               || || :| |  |::|||||| ::|
```

```
Db    957  LNTALDELLKGRTCEVSRYKLIKDTVYKEISDDTELINTIEKNVK----AYLAYIKKNY 1012
                 :|:  |    |  |   |    | | |     ||| |    |         ||  | |
Qy    278  ---ISNLIT--KKISAIKDSGELKAEIEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKK 332

Db   1013  EDTVQDVLTLNEHFNTKQVSNHEPTNFDKSNKSSEELTKAVTDSKTIISK------LKG 1065
                 :|:| |  |      |      | | ||| | |:|| ||||  :   |      ||
Qy    333  AILKTN-----NDKTKGADELEKLFESVKNLSKAAKEM--LTNSVK 371

Db   1066  VIIEVNENTEMMTIESSAKEIEALYNELKNKKTSLNEIYQTSNEVK 1111
                                                          .
RESULT  7

ID   YD86_SCHPO     STANDARD;      PRT;   1957 AA.
AC   Q10411;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   HYPOTHETICAL 222.8 KDA PROTEIN C1F3.06C IN CHROMOSOME I.
GN   SPAC1F3.06C.
OS   Schizosaccharomyces pombe (fission yeast).
OC   Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC   Schizosaccharomycetales; Schizosaccharomycetaceae;
OC   Schizosaccharomyces.
OX   NCBI_TaxID=4896;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=972;
RA   Connor R., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;
RL   Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.
CC   -------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -------------------------------------------------------------
DR   EMBL; Z70690; CAA94624.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE 1957 AA;  222785 MW;  3F480CA06171D99DA CRC64;

Query Match             9.2%;  Score 169;  DB 1;  Length 1957;
Best Local Similarity   23.1%;  Pred. No. 1.2;
Matches  107;  Conservative  73;  Mismatches 156;  Indels 128;  Gaps  20;

Qy      8  KDGNTSANSADE---SVKGPNLTEISKKITDSNAVLLAV-KEVEALLSSIDEI------ 56
              |    |   |     ||  |:| |||   ||    |  :||  |  :|
Db    360  KDSRTSNSQLEEMVELKESNRI-IHSOLTDAESKLSSFEQENKSLKGSIDEYQNNLSSK 418

Qy     57  ---------AAKAIGKKIHONNGLD------------TEYNHNGSLLAGAYAI 88
                     | | |  |                      :|  | || |  :
Db    419  DKMVKQVSSQLEEARSSLAHATGKLAEINSERDFQNKKIKDFEKIEQDLRACLNSSSNEL 478

Qy     89  ---STLIKQKLDGLKNEGLKEKIDAAKKCSETFTNKL-------KEKH------- 126
               :||  |  | | | |     || |  | | ||        |||
Db    479  KEKSALIDKKDQELNN--LREQIKEQKVSESTQSSLQSLQRDILNERKHEYYESQLNE 536

Qy    127  ----------------TDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEM 174
                              |   :| |  | | |    |  :|| | ||  |  ||
Db    537  LKGELQTEISNSEHLSSQLSTLAAEKEAAVATN------NELSESKNSLQTLCNAFQEK 589

Qy    175  LANSVKELTSPVVAESPAMVNNSGKDGNTSANSADESVKGPNLTEISKKITESN-AVVLA 233
           |||:||||||:|: |  |  |    | | |   |   |||  | |  |||  | |  |
Db    590  LAKSVMQL------------KENEQNFSSLDTSFK--KLNESHQELENNHQTITKQ 631

Qy    234  VKEVETLLTSIDELAKAIGKKIKNDVSLDNE----------ADINGSLISGAYLISNLITK 284
           |:|   | |:|||  |  |  |   :|             : |   | |    :| |
Db    632  LKDTSSKLQQL-QLERANFEQKESTLSDENNDLRTKLLKLEESNKSLIKKQEDVDSL-EK 689

Qy    285  KISAIKDSGELKAEIEKAKKCS-----EEFTAKLKGEHTDLGKE-----GVTDDNAKKAI 334
```

```
Db    690  NIQTLKEDLRKSEEALRFSKLEAKNLREVIDNLKGKHETLEAQRNDLHSSLSDAKNTNAI 749
                :|:|    ||    |       ||       |       |||:|  | | | |
Qy    335  LKTNNDKTKGADELEKLFESVKNL---SKAAKEMLTNSVKELTS 375

Db    750  L--SSELTKSSEDVKRLTANVETLTQDSKAMKQSFFSLVNSYQS 791
                 || :|  |  |    | ||  :|  | | | |   ||:  | |

RESULT  8

ID   M6_STRPY       STANDARD;      PRT;    483 AA.
AC   P08089;
DT   01-AUG-1988 (Rel. 08, Created)
DT   01-AUG-1988 (Rel. 08, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   M PROTEIN, SEROTYPE 6 PRECURSOR.
GN   EMM6.
OS   Streptococcus pyogenes.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC   Streptococcus.
OX   NCBI_TaxID=1314;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86111835; PubMed=3511046;
RA   Hollingshead S.K., Fischetti V.F., Scott J.R.;
RT   "Complete nucleotide sequence of type 6 M protein of the group A
RT   streptococcus. Repetitive structure and membrane anchor.";
RL   J. Biol. Chem. 261:1677-1686(1986).
RN   [2]
RP   SEQUENCE OF 43-122 FROM N.A.
RX   MEDLINE=85166224; PubMed=3885219;
RA   Scott J.R., Pulliam W.M., Hollingshead S.K., Fischetti V.A.;
RT   "Relationship of M protein genes in group A streptococci.";
RL   Proc. Natl. Acad. Sci. U.S.A. 82:1822-1826(1985).
CC   -!- FUNCTION: THIS PROTEIN M IS ONE OF THE DIFFERENT ANTIGENIC SEROTYPES
CC       OF PROTEIN M. PROTEIN M IS CLOSELY ASSOCIATED WITH VIRULENCE OF
CC       THE BACTERIUM AND CAN RENDER THE ORGANISM RESISTANT TO
CC       PHAGOCYTOSIS.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. CELL WALL.
CC   -!- SIMILARITY: TO OTHER M PROTEINS.
CC   -!- SIMILARITY: TO OTHER STREPTOCOCCAL AND STAPHYLOCOCCAL PROTEINS
CC       IN THE REGION OF THE MEMBRANE ANCHOR.
CC   -------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -------------------------------------------------------------
DR   EMBL; M11338; AAA26920.1; -.
DR   PIR; A26297; A26297.
DR   InterPro; IPR001899; Gram_pos_anchor.
DR   InterPro; IPR003345; M_repeat.
DR   Pfam; PF00746; Gram_pos_anchor; 1.
DR   Pfam; PF02370; M; 9.
DR   PRINTS; PR00015; GPOSANCHOR.
DR   PROSITE; PS00343; GRAM_POS_ANCHORING; 1.
KW   Virulence; Phagocytosis; Cell wall; Duplication; Repeat; Antigen;
KW   Transmembrane; Coiled coil; Signal.
FT   SIGNAL        1     42
FT   CHAIN        43    483       M PROTEIN, SEROTYPE 6.
FT   DOMAIN       43    457       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    458    477       MEMBRANE ANCHOR.
FT   DOMAIN      478    483       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       69    138       10 X 7 AA TANDEM REPEATS.
FT   DOMAIN      157    269       4.5 X 25 AA TANDEM REPEATS.
FT   DOMAIN      279    347       TWO DIRECTLY REPEATED 27 AMINO ACID
FT                               BLOCKS SEPARATED BY 15 AMINO ACIDS.
FT   DOMAIN      348    411       HYDROPHILIC.
FT   DOMAIN      412    448       GLY/PRO-RICH (CELL WALL-SPANNING).
FT   DOMAIN      449    454       CONSERVED IN GRAM-POSITIVE COCCI SURFACE
```

RESULT 9

ID   M12_STRPY   STANDARD;   PRT;   564 AA.
AC   P19401;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   M PROTEIN, SEROTYPE 12 PRECURSOR (FRAGMENT).
GN   EMM12.
OS   Streptococcus pyogenes.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC   Streptococcus.
OX   NCBI_TaxID=1314;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CS24 / Serotype M12;
RX   MEDLINE=88058777; PubMed=2445730;
RA   Robbins J.C., Spanier J.G., Jones S.J., Cleary P.P.,
RA   Simpson W.J.;
RT   "Streptococcus pyogenes type 12 M protein gene regulation by upstream
RT   sequences.";
RL   J. Bacteriol. 169:5633-5640(1987).
CC   -!- FUNCTION: THIS PROTEIN IS ONE OF THE DIFFERENT ANTIGENIC SEROTYPES
CC   OF PROTEIN M. PROTEIN M IS CLOSELY ASSOCIATED WITH VIRULENCE OF
CC   THE BACTERIUM AND CAN RENDER THE ORGANISM RESISTANT TO
CC   PHAGOCYTOSIS.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. CELL WALL.
CC   -!- SIMILARITY: TO OTHER M PROTEINS.
CC   -!- SIMILARITY: TO OTHER STREPTOCOCCAL AND STAPHYLOCOCCAL PROTEINS
CC   IN THE REGION OF THE MEMBRANE ANCHOR.

Query Match   9.1%; Score 167; DB 1; Length 483;
Best Local Similarity   23.3%; Pred. No. 0.31;
Matches   91; Conservative   57; Mismatches 157; Indels   86; Gaps   14;

Qy   5 NSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGKK 64
          |:|| ::|| | | : | :||| |:| ||: | ||:|| || || || ||:|
Db   73 NNDKLTTENNLLTDQN---KNLITENKNLITDQKNNLITENK------------NL 112

Qy   65 IHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNKLKE 124
          |:|||| |::|| |:| | || ||| || | | :|:| | ||:
Db   113 TDQNKNLITE---NKELKAEENRLTT-------ENKGLTKKLSEAE---EEAANKERE 157

Qy   125 KHTDLGK-EGVTDADAKEAILKTNGTKTKGAEELGKLFESVE--VLSKAAKEM---LAN 177
          || | | ||| | : | | || || |: | | | :|:| ||
Db   158 NKEAIGTKKTLDETVKDIKIAKEQESK------ETIGTLKKTLDETVKDIKIAKEQESKFIG 213

Qy   178 SVKELTSPVVAESPAMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEV 237
          ::|:| | | :||||| | |:|
Db   214 TLKKTLDETVKDIKIAKEQESKEFIGTLKKILDETVK----DKIARE-QKSKQDIGALKQE 268

Qy   238 ETLLTSIDELAKRAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKA 297
          ::||: | : || | |:| | | |: | | |: |: ::| ||:| |
Db   269 LAKKDEGNKVSEASRGLRRDLDASREA------------------KKQVEKDLANLTA 309

Qy   298 EIEKAKKCSEEFTAKLKGEHTDLGKEGVTDDDNAKKAILKTNNDKTKGADELEKLFESVKN 357
          |:|:|| | ||: ||||: ||| | | | || | | : || |:||:|||:|||:|
Db   310 ELDKVKEEKQISDASRQGLRRDLDASREAKKQVEKA-LEEANSKLAALEKLNKELEESKK 368

Qy   358 L-----------------------SKAAKEMLTNSVKEL 373
          | :| |||: |:|:||
Db   369 LTEKEKAEELQAKLEAEAKALKEQLAKQAEEL 399

RESULT 10

ID   USO1_YEAST   STANDARD;   PRT;   1790 AA.
AC   P25386;
DT   01-MAY-1992 (Rel. 22, Created)
DT   01-MAY-1992 (Rel. 22, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   INTRACELLULAR PROTEIN TRANSPORT PROTEIN USO1.
GN   USO1 OR INT1 OR YDL058W.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;

Query Match   9.1%; Score 167; DB 1; Length 564;
Best Local Similarity   22.6%; Pred. No. 0.36;
Matches   97; Conservative   55; Mismatches 161; Indels 116; Gaps   17;

Qy   10 GNTSANSADESVKGPNLTEISKKITDSNAVLLAV-KEVEALLSSIDEIAAKAIGKKIKHON 68
          | ||| || | | | :|| | | :| | ||:| || | :||
Db   115 GKLGIDNADLKAK---ITELEKSVEERKNDVLSQIKKELEEAEKDIQ-----FGREVHAA 165

Qy   69 NGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSEETFTNKLKEKHTD 128
          | |||| : | | | || ||::| | :| : :|:|| | ||:: | ||: | |
Db   166 DLL----RHKQEITAEKENVIS-----KLNG-ELQPLKQKVD----ETDRNLQQE---- 205

Qy   129 LGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVELVLSKAA--KEMLANSVKELTSPV 186
          |||| |:| |:| | :| | || | |: |:|||:| ||:| :||:| |
Db   206 --KQKVLSLEQQLAVTKENAKKDFELAALGHQLADKEYNAKIAELESKLADAKKDFELAA 263

Qy   187 V---NNSGKDG----AESPAMVNNSGKDG-----NTSANSADESVKGP----------- 218
          | | ||| |: | | || || || |||||:|| |
Db   264 LGHQHAHNEYQAKLAEKDGQIKQLEEQKQILDASRKGTARDLEAVRQAKKRATEAELNNLK 323

Qy   219 EISKKITESNAVVLAVK---------------------EVETLLTSIDE---LAKAIGKK 259
          |:| | | | | |: || ||:| || | |||:|
Db   324 AELAKVTEQKQILDASRKGTARDLEAVRKSKKQQVEEAALKQLEEQNKISEASRKGLRRDL 383

Qy   260 IKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAEIEKAKKCSEEFTAKLKG 319
          |::| | | :| |: | :|:| :| | :| | ||:|||||:: ||| | |:|
Db   384 DTSREA-----------------KKQVEKDLANLTAELDKVKEEKQISDASRQGLRRD 424

Qy   320 LGKEGVTDDDNAKKAILKTNNDKTKGADELEKLFESVKN-----------------SKAAKE 364
          | |: |:| | ||:|:| |::| |:|:|:| | |:   :| ||:
Db   425 LDASREAKKQVEKA-LEEANSKLAALEKLNKDLEESKLTEKEKAEELQAKLEAEAKALKE 483

Qy   365 MLTNSVKEL 373
          :|: |:||
Db   484 QLAKQAEEL 492

CC   or send an email to license@isb-sib.ch).
CC
DR   EMBL; M18269; AAA88573.1; -.
DR   PIR; A60115; AAA88573.1; -.
DR   InterPro; IPR001899; Gram_pos_anchor.
DR   InterPro; IPR003445; M_repeat.
DR   Pfam; PF00746; Gram_pos_anchor; 1.
DR   Pfam; PF02370; M; 9.
DR   PROSITE; PS00343; GRAM_POS_ANCHORING; 1.
KW   Virulence; Phagocytosis; Cell wall; Duplication; Repeat; Antigen;
KW   Transmembrane; Coiled coil; Signal.
FT   SIGNAL   1   41
FT   CHAIN   42   >564   M PROTEIN, SEROTYPE 12.
FT   DOMAIN   42   550   EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   551   >564   MEMBRANE ANCHOR.
FT   DOMAIN   44   505   COILED COIL (POTENTIAL).
FT   DOMAIN   505   541   COILED COIL (POTENTIAL).
FT   DOMAIN   542   547   GLY/PRO-RICH (CELL WALL-SPANNING).
FT   NON_TER   564   CONSERVED IN GRAM-POSITIVE COCCI SURFACE
FT   PROTEINS.
SQ   SEQUENCE   564 AA; 62904 MW; 5F15490ACAA77B46 CRC64;

FT
SQ   SEQUENCE   483 AA; 53472 MW; 68F87F28DB53A448 CRC64;

PROTEINS.

Query Match   9.1%; Score 167; DB 1; Length 483;
Best Local Similarity   23.3%; Pred. No. 0.31;
Matches   91; Conservative   57; Mismatches 157; Indels   86; Gaps   14;

RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=X2180-1A;
RX   MEDLINE=91185402; PubMed=2010462;
RA   Nakajima H., Hirata A., Ogawa Y., Yonehara T., Yoda K.,
RA   Yamasaki M.;
RT   "A cytoskeleton-related gene, uso1, is required for intracellular
RT   protein transport in Saccharomyces cerevisiae.";
RL   J. Cell Biol. 113:245-260(1991).
RN   [2]
RP   SEQUENCE OF 1-8 FROM N.A.
RA   Bai Y., Symington L.S.;
RL   Submitted (MAY-1996) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 782-1790 FROM N.A.
RA   Hostetter M.K., Herman D.J., Bendel C.M., McClellan M., Tao N.,
RA   Kendrick K.E.;
RL   Submitted (FEB-1993) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: REQUIRED FOR PROTEIN TRANSPORT FROM THE ER TO THE GOLGI
CC   COMPLEX.
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC. ASSOCIATED WITH INTRACELLULAR
CC   MEMBRANES. PROBABLY PRESENT ON VESICLES OPERATIONAL BETWEEN THE
CC   ER AND THE GOLGI COMPLEX.
CC   -!- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, COMPOSED
CC   OF AN HEPTAPEPTIDE REPEAT PATTERN CHARACTERISTIC OF ALPHA-HELICAL
CC   COILED COILS. MAY FORM FILAMENTOUS STRUCTURES IN THE CELL.
CC   -!- SIMILARITY: BELONGS TO THE VDP/USO1/YBL047C FAMILY.

DR   EMBL; X54378; CAA38253.1; -.
DR   EMBL; L03188; AAB00143.1; -.
DR   EMBL; U53668; AAB66659.1; -.
DR   PIR; A38455; A38455.
DR   HSSP; P80220; 1D1P.
DR   SGD; S0002216; USO1.
DR   InterPro; IPR002017; Spectrin.
KW   Transport; Protein transport; Golgi stack; Cytoskeleton; Coiled coil.
FT   DOMAIN   1   724   GLOBULAR HEAD.
FT   DOMAIN   725   1790   COILED COIL (POTENTIAL).
FT   DOMAIN   465   487   CHARGED (HYPER-HYDROPHILIC).
FT   DOMAIN   991   1790   DISPENSABLE FOR THE PROTEIN FUNCTION.
FT   DOMAIN   1172   1786   ASP/GLU-RICH (ACIDIC).
FT   CONFLICT   847   847   G -> E (IN REF. 2).
FT   CONFLICT   924   924   G -> K (IN REF. 2).
FT   CONFLICT   1253   1253   V -> I (IN REF. 2).
FT   CONFLICT   1319   1319   I -> V (IN REF. 2).
FT   CONFLICT   1461   1461   N -> S (IN REF. 2).
FT   CONFLICT   1581   1581   G -> S (IN REF. 2).
FT   CONFLICT   1600   1600   I -> V (IN REF. 2).
FT   CONFLICT   1661   1661   R -> S (IN REF. 2).
FT   CONFLICT   1772   1772   D -> DEEDDEE (IN REF. 2).
SQ   SEQUENCE   1790 AA; 206424 MW; 6CE2B216F9FD4818 CRC64;

Query Match   8.9%; Score 163.5; DB 1; Length 1790;
Best Local Similarity   21.8%; Pred. No. 1.9;
Matches 102; Conservative   73; Mismatches 184; Indels 109; Gaps   17;

```
Qy     3 CNNSGKDG--------NTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSS 52
           |||  |:      |:     |||:|| :|:: ||:::::
Db   929 CNNLSKEKEHISKELVEYKSRFQOSHDNLV--AKLTEKLKSLANNYKDMQA--ENESLIKA 984

Qy    53 IDEIAAKAIGKKIHQNNGLD--TEYNHNGSLLAGAY---------AISTL------IKOK 95
Db   985 VEESKNESSIQLSNLQNKIDSMSOEKNFOIERGSIEKNIEQLKKTISDLEQTKEEISK 1044
```

RESULT 11
CENF_HUMAN

ID   CENF_HUMAN   STANDARD;   PRT; 3210 AA.
AC   P49454; Q13246; Q13171;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   CENP-F KINETOCHORE PROTEIN (CENTROMERE PROTEIN F) (MITOSIN) (AH
DE   ANTIGEN).
GN   CENPF.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Breast carcinoma;
RX   MEDLINE=95348175; PubMed=7542657;
RA   Liao H., Winkfein R.J., Mack G., Rattner J.B., Yen T.J.;
RT   "CENP-F is a protein of the nuclear matrix that assembles onto
RT   kinetochores at late G2 and is rapidly degraded after mitosis.";
RL   J. Cell Biol. 130:507-518(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95379848; PubMed=7651420;
RA   Zhu X., Mancini M.A., Chang K.-H., Liu C.-Y., Chen C.-F., Shan B.,
RA   Jones D., Yang-Feng T.L., Lee W.-H.;
RT   "Characterization of a novel 350-kilodalton nuclear phosphoprotein
RT   that is specifically involved in mitotic-phase progression.";
RL   Mol. Cell. Biol. 15:5017-5029(1995).
RN   [3]
RP   SEQUENCE OF 2194-3210 FROM N.A.
RX   MEDLINE=95336446; PubMed=7612011;
RA   Li Q., Ke Y., Kapp J.A., Fertig N., Medsger T.A. Jr., Joshi H.C.;
RT   "A novel cell-cycle-dependent 350-kDa nuclear protein: C-terminal
RT   domain sufficient for nuclear localization.";
RL   Biochem. Biophys. Res. Commun. 212:220-228(1995).
RN   [4]
RP   CHARACTERIZATION.
RX   MEDLINE=95370296; PubMed=7642639;
RA   Zhu X., Chang K.-H., He D., Mancini M.A., Brinkley W.R., Lee W.-H.;
RT   "The C terminus of mitosin is essential for its nuclear localization,
RT   centromere/kinetochore targeting, and dimerization.";
RL   J. Biol. Chem. 270:19545-19550(1995).
RN   [5]
RP   CHARACTERIZATION.
RX   MEDLINE=98437347; PubMed=9763420;
RA   Chan G.K.T., Schaar B.T., Yen T.J.;

```
Qy    96 LDGLKNEG-------LKEKIDAAKKCSETFTNKLKE----------------KHDLGKEGVTD 136
          ||||::|       |:||::| :|       |:|
Db  1045 SDSSKDEYESQISLLKEKLEFATTANDENVNKISELTKTREELEAELAAYKNLKNELETK 1104

Qy   137 ADAKEAILK-----------------TNGIKTK-----GAEELGKLFESVEVLSKAA 171
Db  1105 LETSEKALKEVKENEEHLKEEKIQLEKEATETKQQLNSLRANLESLEKEHEDLAAQLKKY 1164

Qy   172 KEMLANSVKELTSPV-----VAESPAMVVNNSGKDGNTSANSADESVKGPNLTEISKKITE 226
Db  1165 EEQIANKEROYNEEISQLNDEITSTQQENESIKKKNDELEGGEVKANKSTSEEQSNLKKSE 1224

Qy   227 SNAVVLAVKEV----ETLLTSIDELAKAIGKKIKINDVSLDNEADHNGSLISGAYLISNLI 282
Db  1225 IDALNLQIKELKKRKNETNEASLLESIKSVESETVKIKELQDECNFREKEVSE------L 1277

Qy   283 TKKISAIKDSG----ELKAEIEKAKKCSEEFTAKLK---GEHTDLGKEGVTDDNAKKAIL 335
Db  1278 EDKLKASEDKNSKYLELQKESEKIKEELDAKTTELKIQLEKITNLSKAKEKSESELSRLK 1337

Qy   336 TKNNDKTKGADE-LEKL--------FESVKNLSKRAAKEMLTNSVKE 372
Db  1338 KTSSEERKNAEEQLEKLKNEIQIKNQAFEKERKLLNEGSSTITQEYSE 1385
```

```
RT   "Characterization of the kinetochore binding domain of CENP-E reveals
RT   interactions with the kinetochore proteins CENP-F and hBUBR1.";
RL   J. Cell Biol. 143:49-63(1998).
CC   -!- FUNCTION: PROBABLY REQUIRED FOR KINETOCHORE FUNCTION, INVOLVED IN
CC       CHROMOSOME SEGREGATION DURING MITOSIS. INTERACTS WITH
CC       RETINOBLASTOMA PROTEIN (RB), CENP-E AND BUB1.
CC   -!- SUBUNIT: HOMO- OR HETERODIMER.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR MATRIX (BUT NOT IN THE NUCLEOLUS),
CC       REORGANIZATION TO THE KINETOCHORE/CENTROMERE (CORONAL SURFACE OF
CC       THE OUTER PLATE) AND THE SPINDLE DURING MITOSIS.
CC   -!- DEVELOPMENTAL STAGE: GRADUALLY ACCUMULATES DURING THE CELL CYCLE.
CC   -!- PTM: HYPERPHOSPHORYLATED DURING MITOSIS.
CC   -------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -------------------------------------------------------------------------
DR   EMBL; U19769; AAA82889.1; -.
DR   EMBL; U30872; AAA82935.1; -.
DR   EMBL; U25725; AAA86889.1; -.
DR   HSSP; P02649; 1LE4.
     MIM; 600236; -.
KW   Chromosomal protein; Nuclear protein; Centromere; Coiled coil;
KW   Mitosis; Phosphorylation; Antigen; Cell cycle; Repeat.
FT   DOMAIN       14    197       COILED COIL (POTENTIAL).
FT   DOMAIN      273    769       COILED COIL (POTENTIAL).
FT   DOMAIN      823   1328       COILED COIL (POTENTIAL).
FT   DOMAIN     1642   1746       COILED COIL (POTENTIAL).
FT   DOMAIN     1862   2987       COILED COIL (POTENTIAL).
FT   DOMAIN     2207   2568       COILED COIL (POTENTIAL).
FT   REPEAT     2207   2386       2 X 177 AA TANDEM REPEATS.
FT   REPEAT     2389   2568       2.
FT   SITE       3015   3032       NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT   CONFLICT     16     16       T -> A (IN REF. 2).
FT   CONFLICT    250    250       L -> Q (IN REF. 2).
FT   CONFLICT    272    272       G -> D (IN REF. 2).
FT   CONFLICT    611    611       G -> D (IN REF. 2).
FT   CONFLICT   1494   1589       MISSING (IN REF. 2).
FT   CONFLICT   1611   1611       V -> A (IN REF. 2).
FT   CONFLICT   1811   1811       V -> A (IN REF. 2).
FT   CONFLICT   2242   2243       ER -> DG (IN REF. 3).
FT   CONFLICT   2335   2335       L -> Q (IN REF. 3).
FT   CONFLICT   2492   2492       D -> N (IN REF. 3).
FT   CONFLICT   2545   2561       ELNERVAALHNDQEACK -> SSMREWQPCIMTKKPVS
FT                                (IN REF. 3).
SQ   SEQUENCE   3210 AA;  367589 MW;  11D83324960E4334 CRC64;
```

```
Query Match             8.8%;  Score 162.5;  DB 1;  Length 3210;
Best Local Similarity   22.0%;  Pred. No. 4;
Matches   84;  Conservative   56;  Mismatches  147;  Indels   95;  Gaps   12;

Qy     44 KEVEALLSSIDEIAAKAIG---KKIHQNNGLDTEVNHNGSLLAGAYAISTLIKQKLDGLK 100
          :||  |:::      :::   ::|            ||| :: |      ||: |   ||
Db   2208 EEVHOLRGIEKLRVRIEADEKKQLHIAERLKERERENDSL------------KDKVENLE 2256

Qy    101 N------EGLKEKIDAAKKCGSETFTNKLKEKHTDLGKEGVTDADAKEAIL 144
                  :|| |:   | | |  |   |  |   |  ||:
Db   2257 RELQMSEENQELVILDAENSKAEVETLKTGIEEMARSLKVFELDL----VTLRSEKENLT 2312

Qy    145 KTNGTKTKGAEELGKLFESVEVL----SKAAKEMLANSVKELTSPVVAES 190
          ||| |  || | |  |   |     |||  ||   |  | |:| || |
Db   2313 KOIQEKQGQLSELDKLLSSFKSLLEEKEQAEIQIKEESKTAVEMLQNQLKELNEAV---- 2368

Qy    191 PAMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLTSIDELAKA 250
          || |  ||  ||  |   :: |  |  ||:  | || ::  |:| |: ::|| | |
Db   2369 ---AALCGQEIMKATEQSLDDPPIEEEHQLRNSIEKLRARLEADEKKQLC 2415

Qy    251 IGKKIKNDVSLDNEADHNGSLISGAYLISNL------ITKKISAIKDSGELKAEIERAK 303
```

```
Db   2416 VLQQLK-------------ESEHHADLLKGR--VENLERELEIARTNQEHAALEAENSKGEVETLK 2466
Qy    304 KCSEEPTAKLKGEHTDL-----GKEGVTDDNAKK------AILKTNNDKTKGADELEKL 351
Db   2467 AKIEGMTQSLRGLEILDVVTIRSEKEDLTNELQKEQERISELEIINSSFENILQEKEQEKV 2526
Qy    352 FESVKNLSKAAKEMLTNSVKEL 373
Db   2527 --QMKEKSSTAPEMLQTQLKEL 2546
```

```
RESULT 12
MYS2_DICDI
ID   MYS2_DICDI     STANDARD;      PRT;   2116 AA.
AC   P08799;
DT   01-NOV-1988 (Rel. 09, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   MYOSIN II HEAVY CHAIN, NON MUSCLE.
GN   MHCA.
OS   Dictyostelium discoideum (Slime mold).
OC   Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
OX   NCBI_TaxID=44689;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87092266; PubMed=3540939;
RA   Warrick H.M., de Lozanne A., Leinwand L.A., Spudich J.A.;
RT   "Conserved protein domains in a myosin heavy chain gene from
RT   Dictyostelium discoideum.";
RL   Proc. Natl. Acad. Sci. U.S.A. 83:9433-9437(1986).
RN   [2]
RP   PHOSPHORYLATION SITES, AND MUTAGENESIS.
RC   STRAIN=AX2;
RX   MEDLINE=90353583; PubMed=2387408;
RA   Lueck-Vielmeter D., Schleicher M., Grabatin B., Wippler J.,
RA   Gerisch G.;
RT   "Replacement of threonine residues by serine and alanine in a
RT   phosphorylatable heavy chain fragment of Dictyostelium myosin II.";
RL   FEBS Lett. 269:239-243(1990).
RN   [3]
RP   PHOSPHORYLATION SITES.
RX   MEDLINE=88112226; PubMed=2828113;
RA   Wagle G., Noegel A., Scheel J., Gerisch G.;
RT   "Phosphorylation of threonine residues on cloned fragments of the
RT   Dictyostelium myosin heavy chain.";
RL   FEBS Lett. 227:71-75(1988).
RN   [4]
RP   X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 1-762.
RX   MEDLINE=95345066; PubMed=7619795;
RA   Fisher A.J., Smith C.A., Thoden J.B., Smith R., Sutoh K., Holden H.M.,
RA   Rayment I.;
RT   "X-ray structures of the myosin motor domain of Dictyostelium
RT   discoideum complexed with MgADP.BeFx and MgADP.AlF4-.";
RL   Biochemistry 34:8960-8972(1995).
RN   [5]
RP   X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 1-762.
RX   MEDLINE=95345067; PubMed=7619796;
RA   Smith C.A., Rayment I.;
RT   "X-ray structure of the magnesium(II)-pyrophosphate complex of the
RT   truncated head of Dictyostelium discoideum myosin to 2.7-A
RT   resolution.";
RL   Biochemistry 34:8973-8981(1995).
RN   [6]
RP   X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 1-762.
RX   MEDLINE=96201189; PubMed=8611530;
RA   Smith C.A., Rayment I.;
RT   "X-ray structure of the magnesium(II) ADP.vanadate complex of the
RT   Dictyostelium discoideum myosin motor domain to 1.9-A resolution.";
RL   Biochemistry 35:5404-5417(1996).
RN   [7]
RP   X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 1-762.
RX   MEDLINE=97452580; PubMed=9305951;
```

RA   Gulick A.M., Bauer C.B., Thoden J.B., Rayment I.;
RT   "X-ray structures of the MgADP, MgATPgammaS, and MgAMPPNP complexes
RT   of the Dictyostelium discoideum myosin motor domain.";
RL   Biochemistry 36:11619-11628(1997).
RN   [8]
RP   X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 1-762.
RX   MEDLINE=98070605; PubMed=9405148;
RA   Bauer C.B., Kuhlman P.A., Bagshaw C.R., Rayment I.;
RT   "X-ray crystal structure and solution fluorescence characterization
RT   of Mg.2'(3')-O-(N-methylanthraniloyl) nucleotides bound to the
RT   Dictyostelium discoideum myosin motor domain.";
RL   J. Mol. Biol. 274:394-407(1997).
CC   -!- FUNCTION: MYOSIN IS A PROTEIN THAT BINDS TO ACTIN & HAS ATPASE
CC   ACTIVITY THAT IS ACTIVATED BY ACTIN.
CC   -!- SUBUNIT: MYOSIN II HEAVY CHAIN IS TWO-HEADED. IT SELF-ASSEMBLES
CC   INTO FILAMENTS. HEXAMER OF 2 HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI
CC   LIGHT CHAIN SUBUNITS (MLC) AND 2 REGULATORY LIGHT CHAIN SUBUNITS
CC   (MLC-2).
CC   -!- SUBCELLULAR LOCATION: HIGHEST CONCENTRATION IN THE POSTERIOR CELL
CC   CORTEX.
CC   -!- DOMAIN: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
CC   MEROMYOSIN (LMM) AND 1 HEAVY MEROMYOSIN (HMM). IT CAN BE FURTHER
CC   SPLIT INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
CC   SUBFRAGMENT (S2).
CC   -!- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CC   CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CC   CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC   -!- PTM: PHOSPHORYLATION INHIBITS THICK FILAMENT FORMATION AND REDUCES
CC   THE ACTIN-ACTIVATED ATPASE ACTIVITY.
CC   -!- MISCELLANEOUS: DICTYOSTELIUM MYOSIN II HAS NO K(2)EDTA ATPASE
CC   ACTIVITY, PERHAPS CORRELATED WITH THE ABSENCE OF A CYS AT THE SH-1
CC   POSITION (688).
CC   -!- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 IQ DOMAIN.
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --
DR   EMBL; M14628; AAA33227.1; -.
DR   PIR; A26655; A26655.
DR   PIR; S00250; S00250.
DR   PDB; 1MMA; 03-DEC-97.
DR   PDB; 1MMD; 17-AUG-96.
DR   PDB; 1MMG; 03-DEC-97.
DR   PDB; 1MMN; 03-DEC-97.
DR   PDB; 1MND; 17-AUG-96.
DR   PDB; 1MNE; 17-AUG-96.
DR   PDB; 1VOM; 23-DEC-96.
DR   PDB; 1LVK; 28-JAN-98.
DR   DictyDb; DD01008; mhcA.
DR   InterPro; IPR000048; IQ.
DR   InterPro; IPR001609; myosin_head.
DR   Pfam; PF00612; IQ; 1.
DR   Pfam; PF00063; myosin_head; 1.
DR   PRINTS; PR00193; MYOSINHEAVY.
DR   ProDom; PD000355; myosin_head; 1.
DR   SMART; SM00015; IQ; 1.
DR   SMART; SM00242; MYSc; 1.
DR   PROSITE; PS50096; IQ; 1.
KW   Myosin; Coiled coil; Actin-binding; ATP-binding; 3D-structure;
KW   Calmodulin-binding; Methylation; Alkylation; Phosphorylation.
FT   DOMAIN   1   761   MYOSIN HEAD-LIKE.
FT   DOMAIN   762   791   IQ.
FT   NP_BIND   179   186   ATP.
FT   DOMAIN   817   2116   COILED COIL (POTENTIAL).
FT   DOMAIN   638   660   ACTIN-BINDING.
FT   DOMAIN   738   752   ACTIN-BINDING.
FT   MOD_RES   130   130   METHYLATION (DI-) (POTENTIAL).

FT   MOD_RES   678   678   ALKYLATION (SH-1).
FT   MOD_RES   1823   1823   PHOSPHORYLATION (BY MHCK).
FT   MOD_RES   1833   1833   PHOSPHORYLATION (BY MHCK).
FT   MOD_RES   2029   2029   PHOSPHORYLATION (BY MHCK).
SQ   SEQUENCE   2116 AA; 243871 MW; 2FC3770BB1EE56A1 CRC64;

Query Match   8.6%; Score 159; DB 1; Length 2116;
Best Local Similarity   21.2%; Pred. No. 3.6;
Matches 98; Conservative 73; Mismatches 153; Indels. 138; Gaps 18;

Qy   9 DGNTSANSADESVKG------PNLTEISKKITDSNAVLLAVKEVEALLSSIDETAAKAIG 62
            :::::::::::::   ::|||||||: : : |||: : ::::|::|::|
Db   964 DGQSDTISRLEKIKDELQKEVEELTESFSEESKDKGVL--EKTRVRLQSELDDIIVR--- 1018

Qy   63 KKIHQNGLDTEYNHNGSLLAGAYAISTLLIKQKLDGLKNEGLKEKIDAAKKCSETFTNKL 122
            ||::::::::: : || : :|:::|::|:|| ||::||:||:|| :::||:| |
Db   1019 ------LDSETKDKSELLRQKKLEEELKQVQEALAAE------TAAKLAQEAANKKL 1064

Qy   123 KEKHFDLGKEGVTDADAKEAILKTNGT------------KTKGA---- 154
            :::|:|| |:| ::: :| | || | |||:|
Db   1065 QGEYTELNEKFNSEVTARSNVEKSKKTLESQLVAVNNELDEEKKNRDALEKKKKALDAML 1124

Qy   155 EELGKLFES--------VEVLSKAAKEMLANSVKELTSPVV--------- 187
            :::|:|| :::|:||::: :: |:|:|:: :
Db   1125 EEMKDQLESTGGGEKKSLYDLKVRQESDMEALRNQISELQSTIAKLEKIKSTLEGEVARLQ 1184

Qy   188 ----AESPAMVNNSGKD-------GNTSANSADESVKGPNLTEISKKITESNAVVLAVKE 236
            ::|:| :::| | : || || ::::::::|::|::::::::|:::||:|:|
Db   1185 GELEAEQLAKSNVEKQKKKVELDLEDKSAQLAEETAAKQALDKLRKKLEQE------LSE 1238

Qy   237 VETLLTSIDELAKAIGKKIK--------NDVSLDNEADHNGSLLSGAYLISNLITK 284
            :|||:|||::|:|:||||| |:::|:: ::|:||:|:||:||:|:|
Db   1239 VQT------QLSEANNKNVNSDSINKHLETSFNNLKLELEAEQKAK------QALEKK 1284

Qy   285 KI----SAIKDSGELKAEIEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNA--KKAILKTNND 340
            || :|::::|:|||:: :::::||:||:|||: : |::|:|:|:::: :: : :||
Db   1285 RLGLESELKHVNEQLEEEKKQKESNEKRKVDLEKEVSEL--KDQIEEEVASKKAVTEAKNK 1343

Qy   341 KTKGADELEKLF-------ESVKNLS--KAAKEMLTNSVKE 372
            ::|::|:|::| ::|::|: :: :::|::::|
Db   1344 KESELDEIKRQYADVVSSRDKSVEQLKTLQAKNEELRNTAEE 1385

RESULT 13
ID   LMB1_MOUSE   STANDARD;   PRT;   1786 AA.
AC   P02469;
DT   21-JUL-1986 (Rel. 01, Created)
DT   01-JUL-1989 (Rel. 11, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   LAMININ BETA-1 CHAIN PRECURSOR (LAMININ B1 CHAIN).
GN   LAMB1-1 OR LAMB-1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87147212; PubMed=3493487;
RA   Sasaki M., Kato S., Kohno K., Martin G.R., Yamada Y.;
RT   "Sequence of the cDNA encoding the laminin B1 chain reveals a
RT   multidomain protein containing cysteine-rich repeats.";
RL   Proc. Natl. Acad. Sci. U.S.A. 84:935-939(1987).
RN   [2]
RP   SEQUENCE OF 1292-1786 FROM N.A.
RX   MEDLINE=85051302; PubMed=6209134;
RA   Barlow D.P., Green N.M., Kurkinen M., Hogan B.L.M.;
RT   "Sequencing of laminin B chain cDNAs reveals C-terminal regions of
RT   coiled-coil alpha-helix.";
RL   EMBO J. 3:2355-2362(1984).
RN   [3]
RP   SEQUENCE OF 165-172; 539-547 AND 712-719.

```
RC   STRAIN=BALB/C; TISSUE=Endothelial cells;
RX   MEDLINE=97363207; PubMed=9219532;
RA   Frieser M., Noeckel H., Pausch F., Roeder C., Hahn A., Deutzmann R.,
RA   Sorokin L.M.;
RT   "Cloning of the mouse laminin alpha 4 cDNA. Expression in a subset of
RT   endothelium.";
RL   Eur. J. Biochem. 246:727-735(1997).
CC   -!- FUNCTION: BINDING TO CELLS VIA A HIGH AFFINITY RECEPTOR, LAMININ
CC       IS THOUGHT TO MEDIATE THE ATTACHMENT, MIGRATION, & ORGANIZATION OF
CC       CELLS INTO TISSUES DURING EMBRYONIC DEVELOPMENT BY INTERACTING
CC       WITH OTHER EXTRACELLULAR MATRIX COMPONENTS.
CC   -!- SUBUNIT: LAMININ IS A COMPLEX GLYCOPROTEIN, CONSISTING OF THREE
CC       DIFFERENT POLYPEPTIDE CHAINS (ALPHA, BETA, GAMMA), WHICH ARE BOUND
CC       TO EACH OTHER BY DISULFIDE BONDS INTO A CROSS-SHAPED MOLECULE
CC       COMPRISING ONE LONG & THREE SHORT ARMS WITH GLOBULES AT EACH END.
CC       THE BETA-1 CHAIN IS A SUBUNIT OF LAMININ-1 (EHS LAMININ), LAMININ-
CC       2 (MEROSIN), AND LAMININ-6 (K-LAMININ).
CC   -!- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC   -!- TISSUE SPECIFICITY: FOUND IN THE BASEMENT MEMBRANES (MAJOR
CC       COMPONENT).
CC   -!- SIMILARITY: CONTAINS 1 LAMININ N-TERMINAL DOMAIN (DOMAIN VI).
CC   -!- SIMILARITY: CONTAINS 12.5 LAMININ EGF-LIKE DOMAINS.
CC   -!- SIMILARITY: CONTAINS 1 LAMININ DOMAIN IV.
CC   --------------------------------------------------------------------------
```

```
DR   EMBL; M15525; AAA39407.1; ALT_INIT.
DR   EMBL; X05212; CAA28839.1; -.
DR   PIR; A26413; MMMSB1.
DR   HSSP; P03069; 1ZIM.
DR   MGD; MGI:96743; Lamb1-1.
DR   InterPro; IPR000561; EGF-like.
DR   InterPro; IPR001886; LamNT.
DR   InterPro; IPR002049; Laminin_EGF.
DR   Pfam; PF00053; laminin_EGF; 13.
DR   Pfam; PF00055; laminin_Nterm; 1.
DR   PRINTS; PR00011; EGFLAMININ.
DR   ProDom; PD002082; LamNT; 1.
DR   SMART; SM00180; LamNT; 1.
DR   SMART; SM00136; EGF_Lam; 11.
DR   PROSITE; PS00022; EGF_1; 9.
DR   PROSITE; PS00186; EGF_2; 2.
DR   PROSITE; PS01248; LAMININ_TYPE_EGF; 11.
DR   PROSITE; PS01186; EGF_2; 2.
KW   Laminin EGF-like domain; Cell adhesion; Repeat; Signal.
KW   Glycoprotein; Basement membrane; Extracellular matrix; Coiled coil;
FT   SIGNAL        1     21
FT   CHAIN        22   1786       LAMININ BETA-1 CHAIN.
FT   DOMAIN       22    270       LAMININ N-TERMINAL (DOMAIN VI).
FT   DOMAIN      271    540       4.5 X LAMININ EGF-LIKE REPEATS (DOMAIN
FT                                V).
FT   DOMAIN      271    334       LAMININ EGF-LIKE 1.
FT   DOMAIN      335    397       LAMININ EGF-LIKE 2.
FT   DOMAIN      398    457       LAMININ EGF-LIKE 3.
FT   DOMAIN      458    509       LAMININ EGF-LIKE 4.
FT   DOMAIN      510    540       LAMININ EGF-LIKE 5 (INCOMPLETE).
FT   DOMAIN      541    772       LAMININ DOMAIN IV.
FT   DOMAIN      773   1178       8 X LAMININ EGF-LIKE REPEATS (DOMAIN
FT                                III).
FT   DOMAIN      773    820       LAMININ EGF-LIKE 6.
FT   DOMAIN      821    866       LAMININ EGF-LIKE 7.
FT   DOMAIN      867    916       LAMININ EGF-LIKE 8.
FT   DOMAIN      917    975       LAMININ EGF-LIKE 9.
FT   DOMAIN      976   1027       LAMININ EGF-LIKE 10.
FT   DOMAIN     1028   1083       LAMININ EGF-LIKE 11.
FT   DOMAIN     1084   1131       LAMININ EGF-LIKE 12.
FT   DOMAIN     1132   1178       LAMININ EGF-LIKE 13.
FT   DOMAIN     1179   1397       DOMAIN II.
FT   DOMAIN     1398   1430       DOMAIN ALPHA.
FT   DOMAIN     1431   1786       DOMAIN I.
FT   DOMAIN     1216   1315       COILED COIL (POTENTIAL).
FT   DOMAIN     1368   1388       COILED COIL (POTENTIAL).
FT   DOMAIN     1778   1786       COILED COIL (POTENTIAL).
FT   DISULFID    271    280       BY SIMILARITY.
FT   DISULFID    273    298       BY SIMILARITY.
FT   DISULFID    300    309       BY SIMILARITY.
FT   DISULFID    312    332       BY SIMILARITY.
FT   DISULFID    335    344       BY SIMILARITY.
FT   DISULFID    337    362       BY SIMILARITY.
FT   DISULFID    365    374       BY SIMILARITY.
FT   DISULFID    377    395       BY SIMILARITY.
FT   DISULFID    398    411       BY SIMILARITY.
FT   DISULFID    400    426       BY SIMILARITY.
FT   DISULFID    428    437       BY SIMILARITY.
FT   DISULFID    440    455       BY SIMILARITY.
FT   DISULFID    458    472       BY SIMILARITY.
FT   DISULFID    460    479       BY SIMILARITY.
FT   DISULFID    481    490       BY SIMILARITY.
FT   DISULFID    493    507       BY SIMILARITY.
FT   DISULFID    773    785       BY SIMILARITY.
FT   DISULFID    775    792       BY SIMILARITY.
FT   DISULFID    794    803       BY SIMILARITY.
FT   DISULFID    806    818       BY SIMILARITY.
FT   DISULFID    821    833       BY SIMILARITY.
FT   DISULFID    823    840       BY SIMILARITY.
FT   DISULFID    842    851       BY SIMILARITY.
FT   DISULFID    854    864       BY SIMILARITY.
FT   DISULFID    867    876       BY SIMILARITY.
FT   DISULFID    869    883       BY SIMILARITY.
FT   DISULFID    886    895       BY SIMILARITY.
FT   DISULFID    898    914       BY SIMILARITY.
FT   DISULFID    917    933       BY SIMILARITY.
FT   DISULFID    919    944       BY SIMILARITY.
FT   DISULFID    946    955       BY SIMILARITY.
FT   DISULFID    958    973       BY SIMILARITY.
FT   DISULFID    976    990       BY SIMILARITY.
FT   DISULFID    978    997       BY SIMILARITY.
FT   DISULFID   1000   1009       BY SIMILARITY.
FT   DISULFID   1012   1025       BY SIMILARITY.
FT   DISULFID   1084   1096       BY SIMILARITY.
FT   DISULFID   1086   1103       BY SIMILARITY.
FT   DISULFID   1105   1114       BY SIMILARITY.
FT   DISULFID   1117   1129       BY SIMILARITY.
FT   DISULFID   1132   1144       BY SIMILARITY.
FT   DISULFID   1134   1151       BY SIMILARITY.
FT   DISULFID   1153   1162       BY SIMILARITY.
FT   DISULFID   1165   1176       BY SIMILARITY.
FT   DISULFID   1179   1179       BY INTERCHAIN (PROBABLE).
FT   DISULFID   1182   1182       BY INTERCHAIN (PROBABLE).
FT   DISULFID   1785   1785       BY INTERCHAIN (PROBABLE).
FT   CARBOHYD    120    120       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    356    356       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    519    519       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    677    677       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD   1041   1041       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD   1195   1195       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD   1279   1279       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD   1336   1336       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD   1343   1343       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD   1487   1487       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD   1533   1533       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD   1542   1542       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD   1643   1643       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CONFLICT   1531   1534       SGNA -> MEMP (IN REF. 2).
FT   CONFLICT   1749   1749       D -> N (IN REF. 2).
SQ   SEQUENCE   1786 AA;  196904 MW;  84667 1B7BF41A474 CRC64;

Query Match          8.6%;  Score 157.5;  DB 1;  Length 1786;
Best Local Similarity   21.8%;  Pred. No. 3.5;
Matches    89;  Conservative  69;  Mismatches 170;  Indels 81;  Gaps  16;
```

```
Qy      3  CNNSGKDG-NTSANSADESVKGPNLTEI------------------SKKITDSNAVLLAVKEVEALLSSI 53
              |:    ||          |    |  |  :: ||  |::  ||| |  :
Db   1406  CSESECCGPNCRTDEGEKKCGGPGCGGLVTVAHSAWQKAMDFDRDVLSALAEVEQLSKMV 1465

Qy     54  DEIAAKAIGKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLIKNEGLKEKIDAAKK 113
             | || |            |       |     |   |||  |  || |   |: |
Db   1466  SEAKVRADEAKQNAQDVL----------------------------------LKTNATKEKVD---K 1495

Qy    114  CSETFTNKLKEKHTDLGKEGVTDADAKEA--------ILKTNGTKTKGAEELGKLFES----VE 165
              |||||||||||||||||||||||||||           |:||  ::| ||  ||       ||
Db   1496  SNEDLRNLIKQIRNFL-TEDSADLDSTEAVANEVLKSGNAST--PQQLQNLTEDIRERVE 1552

Qy    166  VLSKAAKEMLANSVKELTSPVVAESPAMVNNSGKDGNTSANSADESVKGPNLTEIS--KK 223
             |||||||||||||||||||||||||||||||||||||||||||||||||||||:  ||
Db   1553  TLSQVEVILQQSAADIARAELLLEEAKRASKSATDVKVTADMVKEALEEAKAQVAAEKA 1612

Qy    224  ITESNAVVLAVKEVEFLLTSID----------ELAKAIGKKIKNDVSLDNEADHNGSLI 272
             |||||||||||||||:|::|           ||||||:||:||||||||||||||||
Db   1613  IKQADE---DIQGTQNLLTISESETAASEETLTNASQRISKLERNVEELKRKAAQN---- 1665

Qy    273  SGAYLISNLITKKISAIKDSGE------LKAEIEKAKCSEEFTAKLKGEHTDLGKEGVTD 327
             |||||||||||||||||||         |  |:||||   || || ||| ||
Db   1666  SGE---AEYIEKVVYSVKQNADDVKKTLDGELDEKYKYKVESLIAQKTEESADARRKAELL 1722

Qy    328  DNAKRAILKTNNDKTKGADELEKLFE-SVKNLSKAAKEM--LTNSVKEL 373
             |||:|||||||||||||||||:||  :|||||||   |: |:|||||
Db   1723  QNEAKTLLAQANSKLQLLEDLERKYEDNQKYLEDKAQELVRLEGEVRSL 1771
```

```
Query Match             8.5%;   Score 157;  DB 1;  Length 998;
Best Local Similarity  21.1%;   Pred. No. 1.9;
Matches  104;  Conservative  65;  Mismatches 139;  Indels 184;  Gaps 23;

Qy     5  NSG---------KDGNTSANSADESV------KGPNLTEISKKITDSNAVL--LAVKEVEAL 49
             |:||         |||||||||||||          ||:|||   :||||  |   |||:
```

```
Db    424  NAGLTKEKDGNTQIDLINERAITAILNNEKEKQANFITLTKNMVNNNALTPDTKVARVNAV 483

Qy     50  LSSIDEIAAKAIGKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLIKNEGL--KEK 107
             |:|                     ||    |||||||||||||||||||||||||   |||
Db    484  LETI-------------KNNQDTFDIEKSKMLEATVAIT-----------LNSENLTPKQK 520

Qy    108  IDAARKCSETFTNLKEKHTDLGKEGVTDADAKEAILKTN-GTKTKGAE--ELGKLFESV 164
             |||||||||||||||||||||||||||||||||||      |     :|:|| |||||
Db    521  QQMLEKAVDVDLSFKDDTSRAVAIDGITG------AVIKSNLSTKDKGTMLIAVGDKVNAS 575

Qy    165  EVLSKAAKEMLANSV---------KELTSP---------VVAESPAMVNNSGKDG----- 201
             ||||||||||||||          :|||:          ||:||| ||||||||
Db    576  E-LSNAEKQQLLGSVLKKGVETKILSPEQQOLMQQNLDKITAEQTKNDNITEVQGILANP 634

Qy    202  ---NTSANSA------------DESVKGPNLTEISKKITES---------- 227
                |||||:            :||  ||||||:||||
Db    635  AFNTIAKTAAIQKYTTKVLDSPITAEIKGETLESITKIVAESPLNVQDKTDIVKGMGEAI 694

Qy    228  ------NAVVLAVKEVET-----------LLTS-------------IDELAKAI 251
                  |||||||||||            |||             |:||||
Db    695  ASHRTMAPTKKIAAIESVETGVAKSITDLEDKKLMWTKGLVDGIYEDKANPEITSENMKAV 754

Qy    252  GKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAEIEKAKKCSEEFTA 311
             |||||||||||||||||||:|||||||||||||||||||| ||  | |  | |||||
Db    755  SKGVDN----------STAIPEDKQALKDAAS-EAALDRA---TQNFTE 789

Qy    312  KLKGEHTDLGKEGVTDDDNAKKA-----------ILKTNNDKTKGADE--LEKLFESVK 356
             |||||||||||||||||||:|             |||||||||||||   |||||||
Db    790  GLKGQNLDEPKP--RDDIYNKAQDIAYALKNVVTIVLDANPEKREVSEEEVMNKTSSILN 847

Qy    357  NLSKAAKEMLTN 368
             |||||||:|:
Db    848  DISKIAIEKVNN 859
```

```
RP   X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 777-836.
RX   MEDLINE=96419133; PubMed=8805510;
RA   Houdusse A., Cohen C.;
RT   "Structure of the regulatory domain of scallop myosin at 2-A
RT   resolution: implications for regulation.";
RL   Structure 4:21-32(1996).
CC   -!- FUNCTION: MUSCLE CONTRACTION.
CC   -!- FUNCTION: MYOSIN IS A PROTEIN THAT BINDS TO F-ACTIN & HAS ATPASE
CC       ACTIVITY THAT IS ACTIVATED BY F-ACTIN.
CC   -!- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
CC       HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
CC       AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
CC   -!- SUBCELLULAR LOCATION: THICK FILAMENTS OF THE MYOFIBRILS.
CC   -!- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 IQ DOMAIN.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; X55714; CAA39247.1; -.
DR   PIR; S13557; S13557.
DR   PIR; A40997; A40997.
DR   PDB; 1SCM; 30-APR-94.
DR   PDB; 1WDC; 11-JUL-96.
DR   InterPro; IPR000048; IQ.
DR   InterPro; IPR002928; Myosin_tail.
DR   InterPro; IPR002017; Spectrin.
DR   InterPro; IPR001609; myosin_head.
DR   Pfam; PF00612; IQ; 1.
DR   Pfam; PF00063; myosin_head; 1.
DR   Pfam; PF01576; Myosin_tail; 1.
DR   PRINTS; PR00193; MYOSINHEAVY.
DR   ProDom; PD000355; myosin_head; 1.
DR   SMART; SM00015; IQ; 1.
DR   SMART; SM00242; MYSc; 1.
DR   PROSITE; PS50096; IQ; 1.
KW   Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
KW   ATP-binding; Alkylation; Calmodulin-binding; 3D-structure.
FT   DOMAIN        1    777       MYOSIN HEAD-LIKE.
FT   DOMAIN      778    805       IQ.
FT   DOMAIN      836   1938       RODLIKE TAIL (S2 AND LMM DOMAINS).
FT   DOMAIN      836   1938       COILED COIL (POTENTIAL).
FT   NP_BIND     176    183       ATP (BY SIMILARITY).
FT   MOD_RES     693    693       ALKYLATION (SH-1) (BY SIMILARITY).
FT   MOD_RES     703    703       ALKYLATION (SH-2) (BY SIMILARITY).
FT   HELIX       778    821
FT   HELIX       822    823
FT   TURN        825    833
FT   HELIX       834    835
SQ   SEQUENCE   1938 AA;  222821 MW;  A5CCE4127D1A4896 CRC64;
```

```
Query Match         8.5%;
Best Local Similarity   22.2%;
Matches  87;  Conservative  61;  Mismatches 166;  Indels  78;  Gaps  13;

Score 157;  DB 1;  Length 1938;  Pred. No. 4;

Qy    18  DESVKGPNLTFISKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGKKIHQNGLDT----   73
Db   925  DEEDAAADLEGIKKKMEADNA----NLKKDIGDLENTLQKAEQDKAHKDNQISTLGGE  978

Qy    74  ---EYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSEFTNKLKEKHTDLG  130
Db   979  ISQQDEHIGKLNKEKKALEEANKTSDSLQAE--EDKQNHLNKIKAKLEQALDELEDNLE 1036

Qy   131  KEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKELTSPVVAES  190
Db  1037  REKKVRGDVEKA----------KRKVEQDLKSTQENVEDLERVKRE-LEENVRKEAEI---- 1084

Qy   191  PAMVNNSGKDGNTSANSADESVK--GPNLTEISKKI-TESNAVVLAVKEVETLLTSIDEL  247
Db  1085  -SSLNSKLEDEQNLVSQLQRKIKELQARIEELEEELEAERNARKVEKQRAELNRELEEL 1143

Qy   248  AKAIGK----------------------------KIKNDV--SLDNEADHNGSLISGAYIISNLITK  284
Db  1144  GERLDEAGGATSAQIELNKKREAELLKIRRDLEEASLQHEAQ------------ISALRKK 1192

Qy   285  KISAIKDSGELKAEIEKAKKCSEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKG  344
Db  1193  HQDAANEMADQVDQLQKVRSKLEKDKKDLKREMDDLESQ------------MTHNMKNKG 1240

Qy   345  ADE-LEKLFESVKNLSKAAKEMLTNSVKELTS  375
Db  1241  CSEKVMKQFESQMSDLNARLEDSQRSINELQS 1272
```

Search completed: March 18, 2002, 10:11:42
Job time: 960 sec

THIS PAGE BLANK (USPTO)

OM protein - protein search, using sw model

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

Run On: March 18, 2002, 09:55:32 ; Search time 55.5 Seconds
　　　　　　　　　　　　　　　　　　　　(without alignments)
　　　　　　　　　　　　　　　　　　　　153.266 Million cell updates/sec

Title: 1859
Perfect score: 1859
Sequence: 1 MACNNSGKDGNTSANSADES..........KNLSKAAKEMLTNSVKELTS 378

Scoring table: BLOSUM62
Scoring table: Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters:　212252

Minimum DB seq length: 0
Maximum DB seq length: 200000000000

Post-processing: Minimum Match 0%
　　　　　　　　　Maximum Match 100%
　　　　　　　　　Listing first 45 summaries

Database :　Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Match Length	DB ID	Description
1	968.5	52.1	466	4	US-08-235-836C-110	Sequence 110, App
2	945	50.8	210	4	US-08-158-353-3	Sequence 3, Appli
3	938	50.5	210	4	US-08-209-603E-15	Sequence 15, Appli
4	933	50.2	209	4	US-08-196-293-15	Sequence 15, Appli
5	932	50.1	210	4	US-08-235-836C-30	Sequence 30, Appli
6	929	50.0	588	4	US-08-235-836C-122	Sequence 122, App
7	915	49.2	466	4	US-08-235-836C-107	Sequence 107, Appl
8	685	36.8	212	4	US-08-158-353-2	Sequence 2, Appli
9	670	36.0	212	1	US-08-031-295-2	Sequence 2, Appli
10	670	36.0	212	4	US-07-903-580-2	Sequence 2, Appli
11	631	33.9	212	4	US-08-158-353-4	Sequence 4, Appli
12	629	33.8	212	4	US-09-196-293-11	Sequence 11, Appl
13	629	33.8	212	4	US-08-209-603E-11	Sequence 11, Appl
14	626	33.7	212	4	US-08-235-836C-34	Sequence 34, Appl
15	614.5	33.1	209	4	US-08-235-836C-32	Sequence 32, Appl
16	592.5	31.9	207	4	US-08-235-836C-36	Sequence 36, Appli
17	171	9.2	194	4	US-09-364-083-2	Sequence 2, Appli
18	164	8.8	2482	1	US-08-328-254-6	Sequence 6, Appl
19	163	8.8	3248	5	US-08-353-700-1	Sequence 1, Appl
20	163	8.8	3248	2	PCT-US95-16216-1	Sequence 1, Appli
21	155	8.3	1388	2	US-08-685-576-1	Sequence 1, Appli
22	151.5	8.1	1388	2	US-08-685-576-4	Sequence 4, Appli
23	148	8.0	3111	2	US-08-460-309-4	Sequence 4, Appli
24	148	8.0	3111	2	US-08-125-077-4	Sequence 4, Appli
25	146	7.9	688	3	US-09-141-047-8	Sequence 8, Appli
26	145.5	7.8	1196	4	US-08-144-121-4	Sequence 4, Appli
27	145.5	7.8	1196	2	US-08-735-893-4	Sequence 4, Appli
28	144.5	7.8	1164	4	US-08-923-992A-2	Sequence 2, Appli
29	144	7.7	1713	3	US-08-600-982-24	Sequence 24, Appl
30	144	7.7	1713	5	PCT-US94-10261A-24	Sequence 24, Appl
31	143	7.7	630	4	US-08-973-462-9	Sequence 9, Appl
32	142.5	7.7	1164	4	US-08-923-992A-10	Sequence 10, Appl
33	142	7.6	1507	4	US-08-929-329-5	Sequence 5, Appli
34	142	7.6	1786	4	US-08-973-462-8	Sequence 8, Appli
35	139.5	7.5	1663	5	PCT-US93-07261-11	Sequence 11, Appl
36	139.5	7.5	1663	5	PCT-US93-07261-16	Sequence 16, Appl
37	138	7.4	1098	4	US-08-923-992A-8	Sequence 8, Appl
38	137	7.4	641	4	US-08-961-083-160	Sequence 160, Appl
39	136.5	7.3	1038	4	US-09-541-782-4	Sequence 4, Appli
40	135.5	7.3	1104	4	US-08-923-992A-4	Sequence 4, Appli
41	135.5	7.3	1128	4	US-08-923-992A-6	Sequence 6, Appli
42	135	7.2	573	4	US-08-235-836C-112	Sequence 112, App
43	134	7.2	376	6	5180810-1	Patent No. 5180810
44	132	7.1	1561	3	US-08-894-017-23	Sequence 23, Appl
45	131.5	7.1	695	1	US-08-127-499A-23	Sequence 23, Appl

ALIGNMENTS

RESULT 1
US-08-235-836C-110
Sequence 110, Application US/08235836C
Patent No. 6248562

; GENERAL INFORMATION:
; APPLICANT: Dunn, John J.
; APPLICANT: Luft, Benjamin J.
; TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising
; TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor
; NUMBER OF SEQUENCES: 144
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brookhaven National Laboratory.
; STREET:
; CITY: Upton
; STATE: NY
; COUNTRY: USA
; ZIP: 11973
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/235,836C
; FILING DATE: 29-APR-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/148,191
; FILING DATE: 01-11-93
; ATTORNEY/AGENT INFORMATION:
; NAME: Bogosian, Margaret C.
; REGISTRATION NUMBER: 25,324
; REFERENCE/DOCKET NUMBER: BNL93-28A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 282-7338
; TELEFAX: (516) 282-3729
; INFORMATION FOR SEQ ID NO: 110:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 466 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-235-836C-110

Query Match　　52.1%;　Score 968.5;　DB 4;　Length 466;
Best Local Similarity　59.5%;　Pred. No. 5.2e-62;
Matches　229;　Conservative　32;　Mismatches　93;　Indels　31;　Gaps　9;

Qy　1 MACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKA 60

RESULT 2
US-08-158-353-3
Sequence 3, Application US/0815835353
Patent No. 5620862
GENERAL INFORMATION:
  APPLICANT: Padula, Steven J.
  TITLE OF INVENTION: Methods for Diagnosing Early Lyme
  TITLE OF INVENTION: Disease
  NUMBER OF SEQUENCES: 7
  CORRESPONDENCE ADDRESS:
    ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
    STREET: Two Militia Drive
    CITY: Lexington
    STATE: MA
    COUNTRY: USA
    ZIP: 02173
  COMPUTER READABLE FORM:
    MEDIUM TYPE: Floppy disk
    COMPUTER: IBM PC compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: Patentin Release #1.0, Version #1.25
  CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/08/158,353
    FILING DATE:
    CLASSIFICATION: 435
  ATTORNEY/AGENT INFORMATION:
    NAME: Carroll, Alice O.
    REGISTRATION NUMBER: 33,542
    REFERENCE/DOCKET NUMBER: UCT93-05
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: 617-861-6240
    TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 3:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 210 amino acids
    TYPE: amino acid
    STRANDEDNESS: single
    TOPOLOGY: linear
    MOLECULE TYPE: protein
US-08-158-353-3

Query Match   50.8%;
Best Local Similarity   97.9%;
Matches 190; Conservative 3;

Score 945; DB 1; Length 210;
Pred. No. 8.9e-61;
Mismatches 1; Indels 0; Gaps 0;

Qy   17 ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA 76
         ::||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   17 ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA 76

Qy   61 IGKKIHQNNGLLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   77 IGKKIHQNNGLLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN 120

Qy 121 KLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 137 KLKAKHTDLGKEGVTDADAKEAILKTNGTKIKGAEELGKLFESVEVLSKAAKEMLANSVK 196

Qy 181 ELTSPVVAESPKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNA--VVLAVK 238
         |||||||||| | :||: |||||
Db 197 ELTSPVVAESPKKPKPKQ-NVSSLD--EKNSVSVDLPGEMKVLVSKEKKNDGKYDLIATVD 252

Qy 239 EVETLITS------IDELAKAIGKKIKNDVSLDNEADHNGSLLISGAYLI--SNLITKKI 289
         ::| :|:| :|:| ||:   ::||: |:||||:
Db 253 KLELKGTSDKNNGSGVLEGVKADKSKVKLTIS-----DDLGQTTLEVFKEDGKTLVSKKV 307

Qy 290 SAIKDSGELKAEIEAKAKCSEEFTAKLKGEHTDLGKEGVTDNAKKAAILKTNNDKTKGAD 349
         :| |:||| :|:| |:| |||||| :|:|
Db 308 TS-KDKSSTEEKFNEKGEVSEKIITRADG--TRLEYTGIKSDGSGKAKEVLKGTVLEGTL 364

Qy 350 ELEKLFESV-------KNLSKAAK 366
          ||:||:| || :||: :
Db 365 TAEKTTLVVKEGTVTLSKNISKSGE 389

RESULT 3
US-08-209-603E-15
Sequence 15, Application US/0820960315
Patent No. 6248538
GENERAL INFORMATION:
  APPLICANT: FUCHS, RENATE
  APPLICANT: WILSKE, BETTINA
  APPLICANT: PREAC-MURSIC, VERA
  APPLICANT: MOTZ, MANFRED
  APPLICANT: SOUTSCHECK, ERWIN
  TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PROTEINS
  TITLE OF INVENTION: FROM BORRELIA BURGDORFERI
  NUMBER OF SEQUENCES: 15
  CORRESPONDENCE ADDRESS:
    ADDRESSEE: BROOKS HAIDT HAFFNER & DELAHUNTY
    STREET: 99 PARK AVENUE
    CITY: NEW YORK
    STATE: NY
    COUNTRY: USA
    ZIP: 10016
  COMPUTER READABLE FORM:
    MEDIUM TYPE: 3.5" FLOPPY DISC
    COMPUTER: AT&T - IBM COMPATIBLE
    OPERATING SYSTEM: MS-DOS Version 6.2
    SOFTWARE: ASCII
  CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/08/209,603E
    FILING DATE: 10-MAR-1994
    CLASSIFICATION: 436
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: PCT/EP90/02282
    FILING DATE: 21-DEC-1990
  APPLICATION NUMBER:
    FILING DATE: 19-JUN-1992
  ATTORNEY/AGENT INFORMATION:
    NAME: ROBINSON, WILLIAM R.
    REGISTRATION NUMBER: 27,224
    REFERENCE/DOCKET NUMBER: LKR-9217-A
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: (212) 697-3335
    TELEFAX: (212) 557-5635
INFORMATION FOR SEQ ID NO: 15:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 210
    TYPE: AMINO ACID
    TOPOLOGY: LINEAR
  MOLECULE TYPE:
    DESCRIPTION: PROTEIN
  HYPOTHETICAL: N/A
  ANTI-SENSE: N/A
  FRAGMENT TYPE: N/A
  ORIGINAL SOURCE:
    ORGANISM: B. BURGDORFERI
  IMMEDIATE SOURCE:
    LIBRARY: DSM 5662

Qy   1 MACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKA 60
         ::||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db   17 ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA 76

Qy   61 IGKKIHQNNGLLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   77 IGKKIHQNNGLLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN 136

Qy 121 KLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 137 KLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 196

Qy 181 ELTSPVVAESPKKP 194
         ||||||||||||||
Db 197 ELTSPVVAESPKKP 210

```
  POSITION IN GENOME: N/A
  FEATURE:
  IDENTIFICATION METHOD: amino acid analysis
  PUBLICATION INFORMATION: N/A
US-08-209-603E-15

         Query Match
         Best Local Similarity   50.5%;    Score 938; DB 4; Length 210;
         Matches 189; Conservative 97.4%;  Pred. No. 2.8e-60;
                                           Mismatches  2; Indels  0; Gaps  0;

Qy    1  MACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKA  60
         ::||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   17  ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVLEVEALLSSIDEIAAKA  76

Qy   61  IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   77  IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN 136

Qy  121  KLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  137  KLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 196

Qy  181  ELTSPVVAESPKP 194
         |||||||||||||
Db  197  ELTSPVVAESPKP 210
```

```
RESULT 4
US-09-196-293-15
Sequence 15, Application US/09196293
Patent No. 6183755
GENERAL INFORMATION:
  APPLICANT: Fuchs, Renate
  APPLICANT: Motz, Manfred
  APPLICANT: Soitscheck, Erwin
  APPLICANT: Wilske, Bettina
  APPLICANT: Preac-Mursic, Vera
  TITLE OF INVENTION: Active proteins from Borrelia
  TITLE OF INVENTION: burgdorferi
  FILE REFERENCE: 738.001US2
  CURRENT APPLICATION NUMBER: US/09/196,293
  CURRENT FILING DATE: 1998-11-19
  EARLIER APPLICATION NUMBER: US 08/209,603
  EARLIER FILING DATE: 1994-03-10
  EARLIER APPLICATION NUMBER: US 07/862,535
  EARLIER FILING DATE: 1992-06-19
  EARLIER APPLICATION NUMBER: WO PCT/EP90/02282
  EARLIER FILING DATE: 1990-12-21
  EARLIER APPLICATION NUMBER: DE P39 42 728.5
  EARLIER FILING DATE: 1989-12-22
  EARLIER APPLICATION NUMBER: DE P40 18 988.0
  EARLIER FILING DATE: 1990-06-13
  NUMBER OF SEQUENCES: 16
  SOFTWARE: PatentIn for Windows Version 4.0
  SEQ ID NO 15
  LENGTH: 209
  TYPE: PRT
  ORGANISM: Borrelia burgdorferi
US-09-196-293-15

         Query Match
         Best Local Similarity   50.2%;    Score 933; DB 4; Length 209;
         Matches 188; Conservative 97.9%;  Pred. No. 6.4e-60;
                                           Mismatches  1; Indels  0; Gaps  0;

Qy    1  MACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKA  60
         ::||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   17  ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA  76

Qy   61  IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   77  IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN 136

Qy  121  KLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  137  KLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 196

Qy  181  ELTSPVVAESPK 192
         ||||||||||||
Db  197  ELTSPVVAESPK 208
```

```
RESULT 5
US-08-235-836C-30
Sequence 30, Application US/08235836C
Patent No. 6248562
GENERAL INFORMATION:
  APPLICANT: Dunn, John J.
  APPLICANT: Luft, Benjamin J.
  TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising
  TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor
  NUMBER OF SEQUENCES: 144
  CORRESPONDENCE ADDRESS:
    ADDRESSEE: Brookhaven National Laboratory
    STREET:
    CITY: Upton
    STATE: NY
    COUNTRY: USA
    ZIP: 11973
  COMPUTER READABLE FORM:
    MEDIUM TYPE: Floppy disk
    COMPUTER: IBM PC compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: PatentIn Release #1.0, Version #1.25
  CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/08/235,836C
    FILING DATE: 29-APR-1994
    CLASSIFICATION: 435
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: US 08/148,191
    FILING DATE: 01-11-93
  ATTORNEY/AGENT INFORMATION:
    NAME: Bogosian, Margaret C.
    REGISTRATION NUMBER: 25,324
    REFERENCE/DOCKET NUMBER: BNL93-28A
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: (516) 282-7338
    TELEFAX: (516) 282-3729
  INFORMATION FOR SEQ ID NO: 30:
    SEQUENCE CHARACTERISTICS:
      LENGTH: 210 amino acids
      TYPE: amino acid
      TOPOLOGY: linear
      MOLECULE TYPE: protein
US-08-235-836C-30

         Query Match
         Best Local Similarity   50.1%;    Score 932; DB 4; Length 210;
         Matches 188; Conservative 96.9%;  Pred. No. 7.6e-60;
                                           Mismatches  3; Indels  0; Gaps  0;

Qy    1  MACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKA  60
         ::||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db   17  ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA  76

Qy   61  IGKKIHQNNGLDTEYNHNGSLLAGRYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN 120
         ||||||||||||||:|||||||||:||||||||||||||||||||||||||||||||||||
Db   77  IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN 136

Qy  121  KLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  137  KLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 196

Qy  181  ELTSPVVAESPKP 194
         |||||||||||||
Db  197  ELTSPVVAESPKP 210
```

Db 197 ELTSPVVAESPKKP 210

RESULT 6
US-08-235-836C-122
Sequence 122, Application US/0825836C
Patent No. 6248562
GENERAL INFORMATION:
; APPLICANT: Dunn, John J.
; APPLICANT: Luft, Benjamin J.
; TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising
; TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor
; NUMBER OF SEQUENCES: 144
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brookhaven National Laboratory
; STREET:
; CITY: Upton
; STATE: NY
; COUNTRY: USA
; ZIP: 11973
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/235,836C
; FILING DATE: 29-APR-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/148,191
; FILING DATE: 01-11-93
; ATTORNEY/AGENT INFORMATION:
; NAME: Bogosian, Margaret C.
; REGISTRATION NUMBER: 25,324
; REFERENCE/DOCKET NUMBER: BNL93-28A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 282-7338
; TELEFAX: (516) 282-3729
; INFORMATION FOR SEQ ID NO: 122:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 588 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-235-836C-122

Query Match   50.0%; Score 929; DB 4; Length 588;
Best Local Similarity 89.7%; Pred. No. 4.7e-59;
Matches 191; Conservative 6; Mismatches 16; Indels 0; Gaps 0;

Qy   4 NNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGK 63
         |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db 286 NNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGK 345

Qy   64 KIHQNGLDTEYNHNGSLLAGAYAISTLIKQKLDGIKNEGLKEKIDAAKKCSETFTNKLK 123
         ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db 346 KIHQNGLDTEYNHNGSLLAGRYAISTLIKQKLDGIKNEGLKEKIDAAKKCSETFTNKLK 405

Qy 124 EKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKELT 183
         ||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db 406 EKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKELT 465

Qy 184 SPVVAESPKPKPSMVNNSGKDGNTSANSADESVK 216
         ||||||||||||
Db 466 SPVVAESPKPKPGTMAQYNQHMLSNKSASQNVR 498

RESULT 7
US-08-235-8336C-107
; Sequence 107, Application US/08235836C
; Patent No. 6248562

GENERAL INFORMATION:
; APPLICANT: Dunn, John J.
; APPLICANT: Luft, Benjamin J.
; TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising
; TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor
; NUMBER OF SEQUENCES: 144
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brookhaven National Laboratory
; STREET:
; CITY: Upton
; STATE: NY
; COUNTRY: USA
; ZIP: 11973
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/235,836C
; FILING DATE: 29-APR-1994.
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/148,191
; FILING DATE: 01-11-93
; ATTORNEY/AGENT INFORMATION:
; NAME: Bogosian, Margaret C.
; REGISTRATION NUMBER: 25,324
; REFERENCE/DOCKET NUMBER: BNL93-28A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 282-7338
; TELEFAX: (516) 282-3729
; INFORMATION FOR SEQ ID NO: 107:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 466 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-235-836C-107

Query Match   49.2%; Score 915; DB 4; Length 466;
Best Local Similarity 97.4%; Pred. No. 3.5e-58;
Matches 186; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy   4 NNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGK 63
         |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db 276 NNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGK 335

Qy   64 KIHQNGLDTENNHNGSLLAGAYAISTLIKQKLDGIKNEGLKEKIDAAKKCSETFTNKLK 123
         |||||||||| |||||| ||||||||||||||||||||||||||||||||||||||||
Db 336 KIHQNGLDTEYNHNGSLLAGRYAISTLIKQKLDGIKNEGLKEKIDAAKKCSETFTNKLK 395

Qy 124 EKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKELT 183
         ||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db 396 AKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKELT 455

Qy 184 SPVVAESPKKP 194
         |||||||||||
Db 456 SPVVAESPKKP 466

RESULT 8
US-08-158-353-2
; Sequence 2, Application US/08158353
; Patent No. 5620862
GENERAL INFORMATION:
; APPLICANT: Padula, Steven J.
; TITLE OF INVENTION: Methods for Diagnosing Early Lyme
; TITLE OF INVENTION: Disease
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.

; STREET: Two Militia Drive
; CITY: Lexington
; STATE: MA
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/158,353
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; CLASSIFICATION: 435
; NAME: Carroll, Alice O.
; REGISTRATION NUMBER: 33,542
; REFERENCE/DOCKET NUMBER: UC'93-05
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 212 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-158-353-2

RESULT 9
US-08-031-295-2
Sequence 2, Application US/0803129
Patent No. 5530103
; GENERAL INFORMATION:
; APPLICANT: LIVEY, Ian
; APPLICANT: DORNER, Freidrich
; TITLE OF INVENTION: METHOD AND COMPOSITION FOR THE
; TITLE OF INVENTION: PREVENTION OF LYME DISEASE
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25

Query Match 36.8%; Score 685; DB 1; Length 212;
Best Local Similarity 74.1%; Pred. No. 3.8e-42;
Matches 146; Conservative 19; Mismatches 28; Indels 4; Gaps 3;

```
Qy   1 MACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKA 60
         ::||||||||||||||||||||||||||||:|||||||||||||||:||||||:|| ||
Db  17 ISCNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVLLAVKEVETLLASIDEVAKKA 76

Qy  61 IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKN-EGLKEKIDAAKKCSETFT 119
        ||  ||||||||:|||||||||||||||||||||:||||||| ||||||||:||||||||
Db  77 IGNLIAQ-NGLNAGANQNGSLLAGAVVISTLIAEKLDGLKNSEELKEKIEDAKKCNKAFT 135

Qy 120 NKLKEKHTDLG--KEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLAN 177
        :||:| |:|   ||||:||||||||||||||||||||||:|||||||||||||| ||:|
Db 136 DKLKSSHAELGIANGAATDANAKAAILKTNGTKDKGAQELEKLFESVKNLSKAAQETLNN 195

Qy 178 SVKELTSPVVAESPKKP 194
        |||||||||||||:|||
Db 196 SVKELTSPVVAENPKKP 212
```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/031,295
; FILING DATE: 19930312
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/903,580
; FILING DATE: 25-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/824,161
; FILING DATE: 22-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/727,245
; FILING DATE: 11-JUL-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/142 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 212 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-031-295-2

RESULT 10
US-07-903-580-2
Sequence 2, Application US/07903580
Patent No. 6221363
; GENERAL INFORMATION:
; APPLICANT: LIVEY, Ian
; APPLICANT: DORNER, Freidrich
; TITLE OF INVENTION: METHOD AND COMPOSITION FOR THE
; TITLE OF INVENTION: PREVENTION OF LYME DISEASE
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25

Query Match 36.0%; Score 670; DB 1; Length 212;
Best Local Similarity 73.0%; Pred. No. 4.5e-41;
Matches 143; Conservative 15; Mismatches 36; Indels 2; Gaps 2;

```
Qy   1 MACNNSGKDGNT-SANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAK 59
         ::|||||| |: |||| ||||| ||||||||:|||||||||:|||||||||:|||||:|
Db  17 ISCNNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETLVSSIDELATK 76

Qy  60 AIGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEG-LKEKIDAAKKCSETF 118
        |||||| ||||| |   :|||||||||||||||:| | ||||:  ||:||||||||||:|
Db  77 AIGKKIQQNNGLGANADKNGSLLAGAYAISTLITEKLKALKNSGELKAKIEDAKKCSEDF 136

Qy 119 TNKLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANS 178
        |:|| :  | ||:||:|| ||||||||:||||||||||||:||:|||| |||||:| ||S
Db 137 TKKLAAGHAQLGIDGATDNDSKEAILKTNGTKTKGAEELVKLSESVASLSKAAQEASANS 196

Qy 179 VKELTSPVVAESPKKP 194
        ||||||||||||:|||
Db 197 VKELTSPVVAETPKKP 212
```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/903,580
FILING DATE: 19920625
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 435
FILING DATE: 22-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/824,161
FILING DATE: 22-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/727,245
FILING DATE: 11-JUL-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/131 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 212 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-903-580-2

Query Match   36.0%;   Score 670;   DB 4;   Length 212;
Best Local Similarity   73.0%; Pred. No. 4.5e-41;
Matches 143;   Conservative 15;   Mismatches 36;   Indels 2;   Gaps 2;

```
Qy    1 MACNNSGKDGNT-SANSADESVKGPNLTEINKKIITDSNAVLLAVKEVEALLSSIDEIAAK  59
        :||||||| |:  :| |||| ||||||||||:||||||||||:||||||:|:||||:|:|
Db   17 ISCNNSGKGGDSASTNPADESAKGPNLTEISKKIITDSNAFVLAVKEVETLVSSIDELATK  76

Qy   60 AIGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEG-LKEKIDAARKCSETF 118
        |||||| ||||| |:  ||||||||||||||:|||  | |:  ||:|||||:| |||||
Db   77 AIGKKIQQNNGLGANADKNGSLLAGAYAISTLITEKLKRALKNSGELKAKIEDAKKCSEDF 136

Qy  119 TNKLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLPESVEVLSKAAKEMLANS 178
        |:|| | | :||  | ||: ||||||:|||||||||||| ||| ||||||||||:| | |
Db  137 TKKLAAGHAQLGIDGATDNDSKEAILKTNGTKTKGAEELVKLSESVASLSKAAQERSANS 196

Qy  179 VKELTSPVVAESPKKP 194
        ||||||||||| |||
Db  197 VKELTSPVVAETPKKP 212
```

ATTORNEY/AGENT INFORMATION:
NAME: Carroll, Alice O
REGISTRATION NUMBER: 33,542
REFERENCE/DOCKET NUMBER: UCT93-05
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-861-6240
TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 212 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-158-353-4

Query Match   33.9%;   Score 631;   DB 1;   Length 212;
Best Local Similarity   69.9%; Pred. No. 2.8e-38;
Matches 137;   Conservative 15;   Mismatches 42;   Indels 2;   Gaps 2;

```
Qy    1 MACNNSGKDGNT-SANSADESVKGPNLTEINKKIITDSNAVLLAVKEVEALLSSIDEIAAK  59
        :||||||| |:  :| |||| ||||||||||:||||||||||:||||||:|:||||:|:|
Db   17 ISCNNSGKGGDSASTNPADESAKGPNLTEISKKIITDSNAFVLAVKEVETLVSSIDELATK  76

Qy   60 AIGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKN-EGLKEKIDAARKCSETF 118
        ||||  ||||  | |: ||||||||||||||:|| |  | || :||| |:| |||||||
Db   77 AIGKIDNNNGLAALNNQNGSLLAGAYAISTLITEKLSKLKNLEELKTEIARAKKCSEEF 136

Qy  119 TNKLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLPESVEVLSKAAKEMLANS 178
        |||||   : ||  | ||: ||| |||::  |:|||||| | |||||| | |||:|| | |
Db  137 TNKLKSGHADLGKQDATDDHAKAAILKTHATTDKGAKEFKDLPESVEGLLKAAQVALTNS 196

Qy  179 VKELTSPVVAESPKKP 194
        ||||||||||| |||
Db  197 VKELTSPVVAETPKKP 212
```

RESULT 11
US-08-158-353-4
Sequence 4, Application US/08158353
Patent NO. 5620862
GENERAL INFORMATION:
APPLICANT: Padula, Steven J.
TITLE OF INVENTION: Methods for Diagnosing Early Lyme
TITLE OF INVENTION: Disease
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Militia Drive
CITY: Lexington
STATE: MA
COUNTRY: USA
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/158,353
FILING DATE:
CLASSIFICATION: 435

RESULT 12
US-09-196-293-11
Sequence 11, Application US/09196293
Patent NO. 6183755
GENERAL INFORMATION:
APPLICANT: Fuchs, Renate
APPLICANT: Motz, Manfred
APPLICANT: Soutscheck, Erwin
APPLICANT: Wilske, Bettina
APPLICANT: Preac-Mursic, Vera
TITLE OF INVENTION: Active proteins from Borrelia
TITLE OF INVENTION: burgdorferi
FILE REFERENCE: 738.001US2
CURRENT APPLICATION NUMBER: US/09/196,293
EARLIER APPLICATION NUMBER: US/09/196,293
CURRENT FILING DATE: 1998-11-19
EARLIER APPLICATION NUMBER: US 08/209,603
EARLIER FILING DATE: 1994-03-10
EARLIER APPLICATION NUMBER: US 07/862,535
EARLIER FILING DATE: 1992-06-19
EARLIER APPLICATION NUMBER: WO PCT/EP90/02282
EARLIER FILING DATE: 1990-12-21
EARLIER APPLICATION NUMBER: DE P39 42 728.5
EARLIER FILING DATE: 1989-12-22
EARLIER APPLICATION NUMBER: DE P40 18 988.0
EARLIER FILING DATE: 1990-06-13
NUMBER OF SEQ ID NOS: 16
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 11
LENGTH: 212
TYPE: PRT
ORGANISM: Borrelia burgdorferi
US-09-196-293-11

Query Match   33.8%;   Score 629;   DB 4;   Length 212;

Best Local Similarity 70.4%; Pred. No. 3.8e-38; Mismatches 43; Indels 2; Gaps 2;
Matches 138; Conservative 13;

```
Qy   1  MACNNSGKDG-NTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAK  59
        ::||||||||  | ||||| ||||||||||||||||:|||||||||||||:||||||:|
Db  17  ISCNNSGKVGILTSTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETLVLSIDELAKK  76

Qy  60  AIGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKN-EGLKEKIDAAKKCSETF 118
        ||:|| |||| ||  | |||||||||||||||| ||| : |||  |||||:|||||||| |
Db  77  AIGQKIDNNNGLAALNNONGSLLAGAYAISTLITEKLSKLKNLEELKTEIAKAKKCSEEF 136

Qy 119  TNKLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANS 178
        ||||| | || | || |||| |||||| |||| |||||||||||||||||| ||||| ||
Db 137  TNKLKSGHADLGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGLLKAAQVALTNS 196

Qy 179  VKELTSPVVAESPKKP 194
        |||||||||||||||
Db 197  VKELTSPVVAESPKKP 212
```

RESULT 13
US-08-209-603E-11
Sequence 11, Application US/08209603E
Patent No. 6248538
GENERAL INFORMATION:
```
    APPLICANT:  FUCHS, RENATE
    APPLICANT:  WILSKE, BETTINA
    APPLICANT:  PREAC-MURSIC, VERA
    APPLICANT:  MOTZ, MANFRED
    APPLICANT:  SOUTSCHEK, ERWIN
    TITLE OF INVENTION:  IMMUNOLOGICALLY ACTIVE PROTEINS
    TITLE OF INVENTION:  FROM BORRELIA BURGDORFERI
    NUMBER OF SEQUENCES:  15
    CORRESPONDENCE ADDRESS:
      ADDRESSEE:  BROOKS HAIDT HAFFNER & DELAHUNTY
      STREET:  99 PARK AVENUE
      CITY:  NEW YORK
      STATE:  NY
      COUNTRY:  USA
      ZIP:  10016
    COMPUTER READABLE FORM:
      MEDIUM TYPE:  3.5" FLOPPY DISC
      COMPUTER:  AT&T - IBM COMPATIBLE
      OPERATING SYSTEM:  MS-DOS Version 6.2
      SOFTWARE:  ASCII
    CURRENT APPLICATION DATA:
      APPLICATION NUMBER:  US/08/209,603E
      FILING DATE:  10-MAR-1994
      CLASSIFICATION:  436
    PRIOR APPLICATION DATA:
      APPLICATION NUMBER:  PCT/EP90/02282
      FILING DATE:  21-DEC-1990
      APPLICATION NUMBER:  US 07/862,535
      FILING DATE:  19-JUN-1992
    ATTORNEY/AGENT INFORMATION:
      NAME:  ROBINSON, WILLIAM R.
      REGISTRATION NUMBER:  27,224
      REFERENCE/DOCKET NUMBER:  LKR-9217-A
    TELECOMMUNICATION INFORMATION:
      TELEPHONE:  (212) 697-3355
      TELEFAX:  (212) 557-5635
    INFORMATION FOR SEQ ID NO:  11:
      SEQUENCE CHARACTERISTICS:
        LENGTH:  212
        TYPE:  AMINO ACID
        TOPOLOGY:  LINEAR
      MOLECULE TYPE:  PROTEIN
      DESCRIPTION:
      HYPOTHETICAL:  N/A
      ANTI-SENSE:  N/A
      FRAGMENT TYPE:  N/A
      ORIGINAL SOURCE:
        ORGANISM:  B. BURGDORFERI
```

```
    IMMEDIATE SOURCE:
      LIBRARY:  DSM 5662
      POSITION IN GENOME:  N/A
    FEATURE:
      IDENTIFICATION METHOD:  amino acid analysis
      PUBLICATION INFORMATION:  N/A
    ANTI-SENSE:  N/A
    FRAGMENT TYPE:  INTERNAL
    ORIGINAL SOURCE:
US-08-209-603E-11
```

Query Match 33.8%; Score 629; DB 4; Length 212;
Best Local Similarity 70.4%; Pred. No. 3.8e-38;
Matches 138; Conservative 13; Mismatches 43; Indels 2; Gaps 2;

```
Qy   1  MACNNSGKDG-NTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAK  59
        ::||||||||  | ||||| ||||||||||||||||:|||||||||||||:||||||:|
Db  17  ISCNNSGKVGILTSTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETLVLSIDELAKK  76

Qy  60  AIGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKN-EGLKEKIDAAKKCSETF 118
        ||:|| |||| ||  | |||||||||||||||| ||| : |||  |||||:|||||||| |
Db  77  AIGQKIDNNNGLAALNNONGSLLAGAYAISTLITEKLSKLKNLEELKTEIAKAKKCSEEF 136

Qy 119  TNKLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANS 178
        ||||| | || | || |||| |||||| |||| |||||||||||||||||| ||||| ||
Db 137  TNKLKSGHADLGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGLLKAAQVALTNS 196

Qy 179  VKELTSPVVAESPKKP 194
        |||||||||||||||
Db 197  VKELTSPVVAESPKKP 212
```

RESULT 14
US-08-235-836C-34
Sequence 34, Application US/08235836C
Patent No. 6248562
GENERAL INFORMATION:
```
    APPLICANT:  Dunn, John J.
    APPLICANT:  Luft, Benjamin J.
    TITLE OF INVENTION:  No. 6248562el Chimeric Proteins Comprising
    TITLE OF INVENTION:  Borrelia Polypeptides and Uses Therefor
    NUMBER OF SEQUENCES:  144
    CORRESPONDENCE ADDRESS:
      ADDRESSEE:  Brookhaven National Laboratory
      STREET:
      CITY:  Upton
      STATE:  NY
      COUNTRY:  USA
      ZIP:  11973
    COMPUTER READABLE FORM:
      MEDIUM TYPE:  Floppy disk
      COMPUTER:  IBM PC compatible
      OPERATING SYSTEM:  PC-DOS/MS-DOS
      SOFTWARE:  PatentIn Release #1.0, Version #1.25
    CURRENT APPLICATION DATA:
      APPLICATION NUMBER:  US/08/235,836C
      FILING DATE:  29-APR-1994
      CLASSIFICATION:  435
    PRIOR APPLICATION DATA:
      APPLICATION NUMBER:  US 08/148,191
      FILING DATE:  01-11-93
    ATTORNEY/AGENT INFORMATION:
      NAME:  Bogosian, Margaret C.
      REGISTRATION NUMBER:  25,324
      REFERENCE/DOCKET NUMBER:  BNL93-28A
    TELECOMMUNICATION INFORMATION:
      TELEPHONE:  (516) 282-7338
      TELEFAX:  (516) 282-3729
    INFORMATION FOR SEQ ID NO:  34:
      SEQUENCE CHARACTERISTICS:
        LENGTH:  212 amino acids
        TYPE:  amino acid
```

TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-235-836C-34

Query Match
Best Local Similarity 33.7%; Score 626; DB 4; Length 212;
Matches 136; Conservative 16; Mismatches 42; Indels 2; Gaps 2;
Pred. No. 6.3e-38;

```
Qy   1 MACNNSGKDGNT-SANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAK 59
         ::||:||| | ||| | |||:||||||||| |:|||||||| | ||||||:||||||| ||
Db  17 ISCSNSGKGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETLVLSIDELAKK 76

Qy  60 AIGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQRLDGLKN-EGLKEKIDAAKKCSETF 118
         |||:|| |: ||  | ||||||||||||||||  :|| |||||| | |:||||||||||| |
Db  77 AIGQKIDNNNGLAALNNQGSLLAGAYAISTLITEKLSKLKNLEELKTEIAKAKKCSEEF 136

Qy 119 TNKLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANS 178
         ||||| | ||||:| || | |||||||:|  |  ||| | || ||||| ||| |  |
Db 137 TNKLKSGHADLGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGLLKAAQVALINS 196

Qy 179 VKELTSPVVAESPKKP 194
         |||||||||||||||
Db 197 VKELTSPVVAESPKKP 212
```

Matches 133; Conservative 20; Mismatches 39; Indels 3; Gaps 2;

```
Qy   1 MACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKA 60
         ::||:||| |::: : |||:||||||||||||:|||||||| || ||||||:|||||| |
Db  17 ISCNNSG--GDTASTNPDESAKGPNLTVISKKITDSNAFVLAVKEVEALISSIDELANKA 74

Qy  61 IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQRLDGLKN-EGLKEKIDAAKKCSETFT 119
         || || |||:| | ||||||||||||||:| :|| ||| | |:||||||||||| |
Db  75 IGKVIHQNGLNANAGQNGSLLAGAYAISTLITEKLSKLKNSEELNKIEEARKNHSEAFTI 134

Qy 120 NKLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSV 179
         |:|| | || | |  || | |||||||:| | ||| | || ||||| ||| ||||| |
Db 135 NRLKGSHAQLGVAAATDDHAKEAILKSNPTKDKGAKALKDLSESVESLAKAAQEALANSV 194

Qy 180 KELTSPVVAESPKKP 194
         |||| ||||||||||
Db 195 KELTNPVVAESPKKP 209
```

Search completed: March 18, 2002, 09:55:32
Job time: 345 sec

RESULT 15
US-08-235-836C-32
Sequence 32, Application US/0825836C
Patent No. 6248562
GENERAL INFORMATION:
APPLICANT:
APPLICANT: Dunn, John J.
APPLICANT: Luft, Benjamin J.
TITLE OF INVENTION: No. 6248562e1 Chimeric Proteins Comprising
TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor
NUMBER OF SEQUENCES: 144
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brookhaven National Laboratory
STREET:
CITY: Upton
STATE: NY
COUNTRY: USA
ZIP: 11973
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/235,836C
FILING DATE: 29-APR-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/148,191
FILING DATE: 01-11-93
ATTORNEY/AGENT INFORMATION:
NAME: Bogosian, Margaret C.
REGISTRATION NUMBER: 25,324
REFERENCE/DOCKET NUMBER: BNL93-28A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 282-7338
TELEFAX: (516) 282-3729
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 209 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-235-836C-32

Query Match
Best Local Similarity 33.18%; Score 614.5; DB 4; Length 209;
Pred. No. 4.1e-37;

OM protein - protein search, using sw model

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

Run on: March 18, 2002, 10:08:42 ; Search time 621.2 Seconds
                 (without alignments)
                 168.954 Million cell updates/sec

Title: US-09-596-746A-28
Perfect score: 1859
Sequence: 1 MACNNSGKDGNTSANSADES.........KNLSKAAKEMLTNSVKELTS 378

Scoring table: BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched: 3148936 seqs, 27765T034 residues

Total number of hits satisfying chosen parameters: 3148936

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Database :   Pending_Patents_AA_Main:*
   1: /cgn2_6/ptodata/2/paa/PCTUS_COMB.pep:*
   2: /cgn2_6/ptodata/2/paa/US06_COMB.pep:*
   3: /cgn2_6/ptodata/2/paa/US07_COMB.pep:*
   4: /cgn2_6/ptodata/2/paa/US080_COMB.pep:*
   5: /cgn2_6/ptodata/2/paa/US081_COMB.pep:*
   6: /cgn2_6/ptodata/2/paa/US082_COMB.pep:*
   7: /cgn2_6/ptodata/2/paa/US083_COMB.pep:*
   8: /cgn2_6/ptodata/2/paa/US084_COMB.pep:*
   9: /cgn2_6/ptodata/2/paa/US085_COMB.pep:*
  10: /cgn2_6/ptodata/2/paa/US086_COMB.pep:*
  11: /cgn2_6/ptodata/2/paa/US087_COMB.pep:*
  12: /cgn2_6/ptodata/2/paa/US088_COMB.pep:*
  13: /cgn2_6/ptodata/2/paa/US089_COMB.pep:*
  14: /cgn2_6/ptodata/2/paa/US090_COMB.pep:*
  15: /cgn2_6/ptodata/2/paa/US091_COMB.pep:*
  16: /cgn2_6/ptodata/2/paa/US092_COMB.pep:*
  17: /cgn2_6/ptodata/2/paa/US093_COMB.pep:*
  18: /cgn2_6/ptodata/2/paa/US094_COMB.pep:*
  19: /cgn2_6/ptodata/2/paa/US095_COMB.pep:*
  20: /cgn2_6/ptodata/2/paa/US096_COMB.pep:*
  21: /cgn2_6/ptodata/2/paa/US097_COMB.pep:*
  22: /cgn2_6/ptodata/2/paa/US098_COMB.pep:*
  23: /cgn2_6/ptodata/2/paa/US099_COMB.pep:*
  24: /cgn2_6/ptodata/2/paa/US60_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1859	100.0	378	19	US-09-596-746-28	Sequence 28, Appl
2	1859	100.0	378	19	US-09-596-746A-28	Sequence 28, Appl
3	1834	98.7	401	19	US-09-596-746A-80	Sequence 80, Appl
4	1834	98.7	402	19	US-09-596-746A-80	Sequence 80, Appl
5	1810.5	97.4	375	19	US-09-596-746A-24	Sequence 24, Appl
6	1805.5	97.1	374	19	US-09-596-746A-24	Sequence 24, Appl
7	1802.5	97.0	398	19	US-09-596-746-60	Sequence 60, Appl
8	1802.5	97.0	399	19	US-09-596-746A-60	Sequence 60, Appl
9	1614.5	86.8	377	19	US-09-596-746-30	Sequence 30, Appl
10	1614.5	86.8	377	19	US-09-596-746A-30	Sequence 30, Appl
11	1589.5	85.5	400	19	US-09-596-746A-82	Sequence 82, Appl
12	1589.5	85.5	401	19	US-09-596-746A-82	Sequence 82, Appl
13	1566	84.2	374	19	US-09-596-746A-26	Sequence 26, Appl
14	1561	84.0	373	19	US-09-596-746A-26	Sequence 26, Appl
15	1558	83.8	397	19	US-09-596-746A-62	Sequence 62, Appl
16	1558	83.8	398	19	US-09-596-746A-62	Sequence 62, Appl
17	1533.5	82.5	371	19	US-09-596-746A-34	Sequence 34, Appl
18	1528.5	82.2	370	19	US-09-596-746A-34	Sequence 34, Appl
19	1525.5	82.1	395	19	US-09-596-746A-64	Sequence 64, Appl
20	1525.5	82.1	394	19	US-09-596-746A-64	Sequence 64, Appl
21	1501.5	80.8	385	19	US-09-596-746A-56	Sequence 56, Appl
22	1496.5	80.5	384	19	US-09-596-746A-56	Sequence 56, Appl
23	1494.5	80.4	369	19	US-09-596-746A-40	Sequence 40, Appl
24	1493.5	80.3	408	19	US-09-596-746A-78	Sequence 78, Appl
25	1493.5	80.3	409	19	US-09-596-746A-78	Sequence 78, Appl
26	1489.5	80.1	368	19	US-09-596-746A-40	Sequence 40, Appl
27	1486.5	80.0	392	19	US-09-596-746A-72	Sequence 72, Appl
28	1486.5	80.0	393	19	US-09-596-746A-72	Sequence 72, Appl
29	1460.5	78.6	560	19	US-09-596-746A-52	Sequence 52, Appl
30	1460.5	78.6	560	19	US-09-596-746A-52	Sequence 52, Appl
31	1289	69.3	370	19	US-09-596-746A-36	Sequence 36, Appl
32	1284	69.1	369	19	US-09-596-746A-36	Sequence 36, Appl
33	1281	68.9	393	19	US-09-596-746A-66	Sequence 66, Appl
34	1281	68.9	394	19	US-09-596-746A-66	Sequence 66, Appl
35	1250	67.2	368	19	US-09-596-746A-42	Sequence 42, Appl
36	1245	67.0	367	19	US-09-596-746A-42	Sequence 42, Appl
37	1242	66.8	391	19	US-09-596-746A-74	Sequence 74, Appl
38	1242	66.8	392	19	US-09-596-746A-74	Sequence 74, Appl
39	1217	65.5	370	19	US-09-596-746A-32	Sequence 32, Appl
40	1213	65.3	378	19	US-09-596-746A-54	Sequence 54, Appl
41	1213	65.3	378	19	US-09-596-746A-54	Sequence 54, Appl
42	1212	65.2	369	19	US-09-596-746A-32	Sequence 32, Appl
43	1209	65.0	393	19	US-09-596-746A-68	Sequence 68, Appl
44	1209	65.0	394	19	US-09-596-746A-68	Sequence 68, Appl
45	1205	64.8	401	19	US-09-596-746A-70	Sequence 70, Appl

RESULT 1
US-09-596-746-28
Sequence 28, Application US/09596746
GENERAL INFORMATION:
   APPLICANT: Dattwyler, Raymond J.
   APPLICANT: Seinost, Gerald
   APPLICANT: Dykhuizen, Daniel
   APPLICANT: Luft, Benjamin J.
   APPLICANT: Maria J.C. Gomes-Solecki
   TITLE OF INVENTION: Groups of Borrelia burgdorferi and
   TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
   FILE REFERENCE: 2631.1002-001
   CURRENT APPLICATION NUMBER: US/09/596,746
   CURRENT FILING DATE: 2000-06-16
   PRIOR APPLICATION NUMBER: US 60/140,042
   PRIOR FILING DATE: 1999-06-18
   NUMBER OF SEQ ID NOS: 84
   SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 28
   LENGTH: 378
   TYPE: PRT
   ORGANISM: Artificial Sequence
   FEATURE:
   OTHER INFORMATION: OspC Chimera
US-09-596-746-28

ALIGNMENTS
US-09-596-746-28

Query Match 100.0%; Score 1859; DB 19; Length 378;
Best Local Similarity 100.0%; Pred. No. 4.5e-123;
Matches 378; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 MACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKA 60

Left column (RESULT 2)

```
Db      1  MACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKA
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     61  IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN

Db     61  IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    121  KLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK  60

Db    121  KLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    181  ELTSPVVAESPKKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEV 120

Db    181  ELTSPVVAESPKKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEV 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    241  ETLLTSIDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKA 180

Db    241  ETLLTSIDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKA 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    301  EIEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKAILKTNNDKTKGADELEKLFESVKN  240

Db    301  EIEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKAILKTNNDKTKGADELEKLFESVKN 360
           |||||||||||||||||||
Qy    361  LSKAAKEMLINSVKELTS 378

Db    361  LSKAAKEMLINSVKELTS 378
```

Query Match
Best Local Similarity 100.0%;
Matches 378; Conservative 100.0%; Mismatches 0; Indels 0; Gaps 0;
                       Score 1859; DB 19; Length 378;
                       Pred. No. 4.5e-123;

Right column (RESULT 3)

```
Db    181  ELTSPVVAESPKKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEV
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    241  ETLLTSIDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKA

Db    241  ETLLTSIDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKA 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    301  EIEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKAILKTNNDKTKGADELEKLFESVKN 300

Db    301  EIEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKAILKTNNDKTKGADELEKLFESVKN 300
           |||||||||||||||||||
Qy    361  LSKAAKEMLINSVKELTS 378

Db    361  LSKAAKEMLINSVKELTS 378
```

Query Match
Best Local Similarity 98.7%;
Matches 373; Conservative 98.9%; Mismatches 2; Indels 0; Gaps 0;
                       Score 1834; DB 19; Length 401;
                       Pred. No. 2.9e-121;

```
Qy      2  ACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKAI
           :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     25  SCNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKAI  61

Qy     62  GKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNK  61
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     85  GKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNK  84

Qy    122  LKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE  84
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    145  LKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 144

Qy    182  LTSPVVAESPKKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVE 144
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    205  LTSPVVAESPKKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVE 204

Qy    242  TLLTSIDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAE 204
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    265  TLLTSIDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAE 264

Qy    302  IEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKAILKTNNDKTKGADELEKLFESVKNL 264
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    325  IEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKAILKTNNDKTKGADELEKLFESVKNL 324

Qy    362  SKAAKEMLINSVKELTS 301
           |||||||||||||||||
Db    385  SKAAKEMLINSVKELTS 384
```

RESULT 4
US-09-596-746A-80
Sequence 80, Application US/0959674A
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seinost, Gerald
; APPLICANT: Dykhuizen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746A
; CURRENT FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 80
; LENGTH: 402
; TYPE: PRT
; ORGANISM: ospc Chimera
US-09-596-746A-80

Query Match   98.7%;   Score 1834;   DB 19;   Length 402;
Best Local Similarity   98.9%;   Pred. No. 2.9e-121;
Matches 373;   Conservative   2;   Mismatches   2;   Indels   0;   Gaps   0;

```
Qy    2 ACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKAI   61
          :|||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db   26 SCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAI   85

Qy   62 GKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNK  121
         |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db   86 GKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNK  145

Qy  122 LKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE  181
         |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db  146 LKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE  205

Qy  182 LTSPVVAESPKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVE   241
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  206 LTSPVVAESPKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVE   265

Qy  242 TLLTSIDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAE  301
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  266 TLLTSIDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAE  325

Qy  302 IEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNL  361
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  326 IEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNL  385

Qy  362 SKAAKEMLTNSVKELTS  378
         |||||||||||||||||
Db  386 SKAAKEMLTNSVKELTS  402
```

RESULT 5
US-09-596-746A-24
Sequence 24, Application US/0959674A
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seinost, Gerald
; APPLICANT: Dykhuizen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746A
; CURRENT FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042

; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: OspC Chimera
US-09-596-746A-24

Query Match   97.4%;   Score 1810.5;   DB 19;   Length 375;
Best Local Similarity   98.1%;   Pred. No. 1.2e-119;
Matches 371;   Conservative   2;   Mismatches   2;   Indels   3;   Gaps   1;

```
Qy    1 MACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKA   60
         ||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db    1 MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA   60

Qy   61 IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN  120
         ||||||||||||| || |||||||||||||||||||||||||||||||||||||||||||
Db   61 IGKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN  120

Qy  121 KLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK  180
         ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db  121 KLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK  180

Qy  181 ELTSPVVAESPKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEV   240
         ||||||||||   :|||||||||||||||||||||||||||||||||||||||||||||
Db  181 ELTSPVVAES---PAMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEV  237

Qy  241 ETLLTSIDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKA  300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  238 ETLLTSIDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKA  297

Qy  301 EIEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKN  360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  298 EIEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKN  357

Qy  361 LSKAAKEMLTNSVKELTS  378
         ||||||||||||||||||
Db  358 LSKAAKEMLTNSVKELTS  375
```

RESULT 6
US-09-596-746-24
Sequence 24, Application US/0959674A
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seinost, Gerald
; APPLICANT: Dykhuizen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746
; CURRENT FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: OspC Chimera
US-09-596-746-24

Query Match   97.1%;   Score 1805.5;   DB 19;   Length 374;

```
Best Local Similarity  98.1%;  Conservative 98.1%;  Pred. No. 2.7e-119;
Matches 370;  Conservative 2;  Mismatches 2;  Indels 3;  Gaps 1;

Qy   2  ACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKAI  61
        |||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db   1  ACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAI  60

Qy  62  GKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSEFTFNK  121
        |||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db  61  GKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSEFTFNK  120

Qy 122  LKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE  181
        |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db 121  LKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE  180

Qy 182  LTSPVVAESPKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVE  241
        |||||||||:  ||||||||||||||||||||||||||||||||||||||||||||||||
Db 181  LTSPVVAES---PAMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVE  237

Qy 242  TLLTSIDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAE  301
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 238  TLLTSIDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAE  301

Qy 302  IEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNL  361
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 298  IEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNL  361

Qy 362  SKAAKEMLTNSVKELTS  378
        |||||||||||||||||
Db 358  SKAAKEMLTNSVKELTS  374
```

RESULT 7
US-09-596-746-60
Sequence 60, Application US/0959676746
GENERAL INFORMATION:
;   APPLICANT: Dattwyler, Raymond J.
;   APPLICANT: Seinost, Gerald
;   APPLICANT: Dykhuizen, Daniel
;   APPLICANT: Luft, Benjamin J.
;   APPLICANT: Maria J.C. Gomes-Solecki
;   TITLE OF INVENTION: Groups of Borrelia
;   TITLE OF INVENTION: burgdorferi and
;   TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
;   FILE REFERENCE: 2631.1002-001
;   CURRENT APPLICATION NUMBER: US/09/596,746
;   CURRENT FILING DATE: 2000-06-16
;   PRIOR APPLICATION NUMBER: US 60/140,042
;   PRIOR FILING DATE: 1999-06-18
;   NUMBER OF SEQ ID NOS: 84
;   SOFTWARE: FastSEQ for Windows Version 4.0
;   SEQ ID NO 60
;   LENGTH: 398
;   TYPE: PRT
;   ORGANISM: ospC Chimera
US-09-596-746-60

Query Match   97.0%;   Score 1802.5;   DB 19;   Length 398;
Best Local Similarity 97.9%; Pred. No. 4.8e-119;
Matches 369; Conservative 3; Mismatches 2; Indels 3; Gaps 1;

```
Qy   2  ACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKAI  61
        :|||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db  25  SCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAI  84

Qy  62  GKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSEFTFNK  121
        |||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db  85  GKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSEFTFNK  144

Qy 122  LKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE  181
        |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db 145  LKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE  204

Qy 182  LTSPVVAESPKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVE  241
        |||||||||:  ||||||||||||||||||||||||||||||||||||||||||||||||
Db 205  LTSPVVAES---PAMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVE  261

Qy 242  TLLTSIDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAE  301
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 262  TLLTSIDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAE  301

Qy 302  IEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNL  361
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 322  IEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNL  381

Qy 362  SKAAKEMLTNSVKELTS  378
        |||||||||||||||||
Db 382  SKAAKEMLTNSVKELTS  398
```

RESULT 8
US-09-596-746A-60
Sequence 60, Application US/0959676746A
GENERAL INFORMATION:
;   APPLICANT: Dattwyler, Raymond J.
;   APPLICANT: Seinost, Gerald
;   APPLICANT: Dykhuizen, Daniel
;   APPLICANT: Luft, Benjamin J.
;   APPLICANT: Maria J.C. Gomes-Solecki
;   TITLE OF INVENTION: Groups of Borrelia burgdorferi and
;   TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
;   FILE REFERENCE: 2631.1002-001
;   CURRENT APPLICATION NUMBER: US/09/596,746A
;   CURRENT FILING DATE: 2000-06-19
;   PRIOR APPLICATION NUMBER: US 60/140,042
;   PRIOR FILING DATE: 1999-06-18
;   NUMBER OF SEQ ID NOS: 84
;   SOFTWARE: FastSEQ for Windows Version 4.0
;   SEQ ID NO 60
;   LENGTH: 399
;   TYPE: PRT
;   ORGANISM: ospC Chimera
US-09-596-746A-60

Query Match   97.0%;   Score 1802.5;   DB 19;   Length 399;
Best Local Similarity 97.9%; Pred. No. 4.8e-119;
Matches 369; Conservative 3; Mismatches 2; Indels 3; Gaps 1;

```
Qy   2  ACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKAI  61
        :|||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db  26  SCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAI  85

Qy  62  GKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSEFTFNK  121
        |||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db  86  GKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSEFTFNK  145

Qy 122  LKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE  181
        |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db 146  LKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE  205

Qy 182  LTSPVVAESPKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVE  241
        |||||||||:  ||||||||||||||||||||||||||||||||||||||||||||||||
Db 206  LTSPVVAES---PAMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVE  262

Qy 242  TLLTSIDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAE  301
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 263  TLLTSIDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAE  322

Qy 302  IEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNL  361
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 323  IEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNL  382

Qy 362  SKAAKEMLTNSVKELTS  378
        |||||||||||||||||
Db 383  SKAAKEMLTNSVKELTS  399
```

RESULT 9
US-09-596-746-30
Sequence 30, Application US/0959674 6
GENERAL INFORMATION:
APPLICANT: Dattwyler, Raymond J.
APPLICANT: Seinost, Gerald
APPLICANT: Dykhuizen, Daniel
APPLICANT: Luft, Benjamin J.
APPLICANT: Maria J.C. Gomes-Solecki
TITLE OF INVENTION: Groups of Borrelia burgdorferi and
TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
FILE REFERENCE: 2631.1002-001
CURRENT APPLICATION NUMBER: US/09,596,746
CURRENT FILING DATE: 2000-06-16
PRIOR APPLICATION NUMBER: US 60/140,042
PRIOR FILING DATE: 1999-06-18
NUMBER OF SEQ ID NOS: 84
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 30
LENGTH: 377
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: OspC Chimera
US-09-596-746-30

Query Match   86.8%; Score 1614.5; DB 19; Length 377;
Best Local Similarity 87.8%; Pred. No. 9e-106;
Matches 331; Conservative 20; Mismatches 25; Indels   1; Gaps   1;

Qy   1 MACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKA 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKA 60

Qy   61 IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN 120

Qy 121 KLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 KLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 180

Qy 181 ELTSPVVAESPKKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEV 240
          |||||||||||||||||||||||||||||||||||:::|:|:::.:|:|:.|||
Db 181 ELTSPVVAESPKKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEI 240

Qy 241 ETLLTSIDELA-KAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELK 299
          |||:|||||| :||.||:|||||:||||||.|||:|:||:|||:|:|:||.|:|
Db 241 ETLLASIDELATKAIGKKIQONGGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLK 300

Qy 300 AEIEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVK 359
          :|| ||||||:|.|| |||:.||||||||||.| || ||||:|:||||.| ||||::|:
Db 301 EKIENAKKCSEDFTKKLEGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVE 360

Qy 360 NLSKAAKEMLTNSVKEL 376
          |:|||||||.|||||
Db 361 NLAKAAKEMLANSVKEL 377

RESULT 10
US-09-596-746A-30
Sequence 30, Application US/0959674 6A
GENERAL INFORMATION:
APPLICANT: Dattwyler, Raymond J.
APPLICANT: Seinost, Gerald
APPLICANT: Dykhuizen, Daniel
APPLICANT: Luft, Benjamin J.
APPLICANT: Maria J.C. Gomes-Solecki
TITLE OF INVENTION: Groups of Borrelia burgdorferi and
TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans

---

FILE REFERENCE: 2631.1002-001
CURRENT APPLICATION NUMBER: US/09,596,746A
CURRENT FILING DATE: 2000-06-19
PRIOR APPLICATION NUMBER: US 60/140,042
PRIOR FILING DATE: 1999-06-18
NUMBER OF SEQ ID NOS: 84
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 30
LENGTH: 377
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: OspC Chimera
US-09-596-746A-30

Query Match   86.8%; Score 1614.5; DB 19; Length 377;
Best Local Similarity 87.8%; Pred. No. 9e-106;
Matches · 331; Conservative 20; Mismatches 25; Indels   1; Gaps   1;

Qy   1 MACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKA 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKA 60

Qy   61 IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN 120

Qy 121 KLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 KLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 180

Qy 181 ELTSPVVAESPKKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEV 240
          |||||||||||||||||||||||||||||||||||:::|:|:::.:|:|:.|||
Db 181 ELTSPVVAESPKKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEI 240

Qy 241 ETLLTSIDELA-KAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELK 299
          |||:|||||| :||.||:|||||:||||||.|||:|:||:|||:|:|:||.|:|
Db 241 ETLLASIDELATKAIGKKIQONGGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLK 300

Qy 300 AEIEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVK 359
          :|| ||||||:|.|| |||:.||||||||||.| || ||||:|:||||.| ||||::|:
Db 301 EKIENAKKCSEDFTKKLEGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVE 360

Qy 360 NLSKAAKEMLTNSVKEL 376
          |:|||||||.|||||
Db 361 NLAKAAKEMLANSVKEL 377

RESULT 11
US-09-596-746-82
Sequence 82, Application US/0959674 6
GENERAL INFORMATION:
APPLICANT: Dattwyler, Raymond J.
APPLICANT: Seinost, Gerald
APPLICANT: Dykhuizen, Daniel
APPLICANT: Luft, Benjamin J.
APPLICANT: Maria J.C. Gomes-Solecki
TITLE OF INVENTION: Groups of Borrelia burgdorferi and
TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
FILE REFERENCE: 2631.1002-001
CURRENT APPLICATION NUMBER: US/09,596,746
CURRENT FILING DATE: 2000-06-16
PRIOR APPLICATION NUMBER: US 60/140,042
PRIOR FILING DATE: 1999-06-18
NUMBER OF SEQ ID NOS: 84
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 82
LENGTH: 400
TYPE: PRT
ORGANISM: OspC Chimera
US-09-596-746-82

RESULT 12
US-09-596-746A-82
Sequence 82, Application US/0959674A6A
GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seinost, Gerald
; APPLICANT: Dykhuizen, Danial
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746A
; CURRENT FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 82
; LENGTH: 401
; TYPE: PRT
; ORGANISM: OspC Chimera

US-09-596-746A-82

Query Match   85.8%;   Score 1589.5;   DB 19;   Length 400;
Best Local Similarity   86.7%;   Pred. No. 5.7e-104;
Matches 326;   Conservative   22;   Mismatches   27;   Indels   1;   Gaps   1;

Qy   2 ACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKAI   61
          ||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db   25 SCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAI   84

Qy   62 GKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNK 121
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   85 GKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNK 144

Qy   122 LKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 181
          ||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db   145 LKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 204

Qy   182 LTSPVVAESPKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVE   241
          ||||||| ||||||||||||||||||||||||||||||||||||||| ||||| |:|
Db   205 LTSPVVAESPKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIE   264

Qy   242 TLLTSIDELA-KAIGKKIKNDVSLDNEADHNGSLLSGAYLISNLIITKKISAIKDSGELKA 300
          ||| |||||| ||||||||||||||||||||||||| |||| |:| |||:||||||||||
Db   265 TLLASIDELATKAIGKKIQONGGLAVEAGHNGTLLAGAYTISKLIITQKLDGLKNSEKLKE 324

Qy   301 EIEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKN 360
          :|| |||||||||| | |||||||||||:||||| | |:||| ::|:|:||||||||| :|
Db   325 KIENAKKCSEDFTKKLEGEHAQLGIENVTDENAKKAILTDAAKDKGAAELELKFKAVEN 384

Qy   361 LSKAAKEMLTNSVKEL 376
          :|:||||||:|||||
Db   385 LAKAAKEMLANSVKEL 400

RESULT 13
US-09-596-746A-26
Sequence 26, Application US/0959674A6A
GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seinost, Gerald
; APPLICANT: Dykhuizen, Danial
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746A
; CURRENT FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 26
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: OspC Chimera

US-09-596-746A-26

Query Match   84.2%;   Score 1566;   DB 19;   Length 374;
Best Local Similarity   85.9%;   Pred. No. 2.4e-102;
Matches 324;   Conservative   22;   Mismatches   27;   Indels   4;   Gaps   2;

Qy   1 MACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKA   60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA   60

Qy   61 IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 IGKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN 120

Qy   121 KLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 180
          |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db   121 KLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 180

Qy   181 ELTSPVVAESPKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEV   240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 ELTSPVVAES---PAMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEI   237

Qy   241 ETLLTSIDELA-KAIGKKIKNDVSLDNEADHNGSLLSGAYLISNLIITKKISAIKDSGELK 299
          ||| |||||| ||||||||||||||||||||||||| ||||| |:| |||:||||||||:
Db   238 ETLLASIDELATKAIGKKIQONGGLAVEAGHNGTLLAGAYTISKLIITQKLDGLKNSEKLK 297

Qy   300 AEIEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVK 359
          :|| |||||||||| | |||||||||||:||||| | |:||| ::|:|:||||||||| |
Db   298 EKIENAKKCSEDFTKKLEGEHAQLGIENVTDENAKKAILTDAAKDKGAAELELKFKAVE 357

Qy   360 LSKAAKEMLTNSVKEL 376
          :|:||||||:|||||
Db   386 LAKAAKEMLANSVKEL 401

RESULT 14
US-09-596-746-26
Sequence 26, Application US/09596746
GENERAL INFORMATION:
;   APPLICANT: Dattwyler, Raymond J.
;   APPLICANT: Selnost, Gerald
;   APPLICANT: Dykhuizen, Danial
;   APPLICANT: Luft, Benjamin J.
;   APPLICANT: Maria J.C. Gomes-Solecki
;   TITLE OF INVENTION: Groups of Borrelia burgdorferi and
;   TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
;   FILE REFERENCE: 2631.1002-001
;   CURRENT APPLICATION NUMBER: US/09/596,746
;   CURRENT FILING DATE: 2000-06-16
;   PRIOR APPLICATION NUMBER: US 60/140,042
;   PRIOR FILING DATE: 1999-06-18
;   NUMBER OF SEQ ID NOS: 84
;   SOFTWARE: FastSEQ for Windows Version 4.0
;   SEQ ID NO 26
;   LENGTH: 373
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: OspC Chimera
US-09-596-746-26

Query Match   84.0%; Score 1561; DB 19; Length 373;
Best Local Similarity   85.9%; Pred. No. 5.4e-102;
Matches 323; Conservative 22; Mismatches 27; Indels   4; Gaps   2;

Qy   2 ACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKAI 61
           |:||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Db   1 ACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAI 60

Qy   62 GKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNK 121
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 GKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNK 120

Qy   122 LKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 181
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 LKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 180

Qy   182 LTSPVVAESPKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVE 241
           |:||||||||| |:||||||||||||||||||||||||||||||||||||:|
Db   181 LTSPVVAES---PAMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIE 237

Qy   242 TLLTSIDELA-KAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKA 300
           ||||||||| |:||||||||::   : ||||||:|||||||||||::|||:|
Db   238 TLLASIDELATKAIGKKIQQNGLAVEAGHNGTLLAGAYTISKLITOKLDGLKNSEKLKE 297

Qy   301 EIEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKN 360
           :|| |||||||| | | ||| ||||||||||:|| | |||||||||||||||:|
Db   298 KIENAKKCSEDFTKKLEGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVEN 357

Qy   361 LSKAAKEMLTNSVKEL 376
           |:||||||| |||||
Db   358 LAKAAKEMLANSVKEL 373

Qy   360 NLSKAAKEMLTNSVKEL 376
           |:|||||||| |||||
Db   358 NLAKAAKEMLANSVKEL 374

RESULT 15
US-09-596-746-62
Sequence 62, Application US/09596746
GENERAL INFORMATION:
;   APPLICANT: Dattwyler, Raymond J.
;   APPLICANT: Selnost, Gerald
;   APPLICANT: Dykhuizen, Danial
;   APPLICANT: Luft, Benjamin J.
;   APPLICANT: Maria J.C. Gomes-Solecki
;   TITLE OF INVENTION: Groups of Borrelia burgdorferi and
;   TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
;   FILE REFERENCE: 2631.1002-001
;   CURRENT APPLICATION NUMBER: US/09/596,746
;   CURRENT FILING DATE: 2000-06-16
;   PRIOR APPLICATION NUMBER: US 60/140,042
;   PRIOR FILING DATE: 1999-06-18
;   NUMBER OF SEQ ID NOS: 84
;   SOFTWARE: FastSEQ for Windows Version 4.0
;   SEQ ID NO 62
;   LENGTH: 397
;   TYPE: PRT
;   ORGANISM: OspC Chimera
US-09-596-746-62

Query Match   83.8%; Score 1558; DB 19; Length 397;
Best Local Similarity   85.6%; Pred. No. 9.6e-102;
Matches 322; Conservative 23; Mismatches 27; Indels   4; Gaps   2;

Qy   2 ACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKAI 61
           |:||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Db   25 SCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAI 84

Qy   62 GKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNK 121
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   85 GKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNK 144

Qy   122 LKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 181
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   145 LKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 204

Qy   182 LTSPVVAESPKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVE 241
           |:||||||| |:|||||||||||||||||||||||||||||||||||||:|
Db   205 LTSPVVAES---PAMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIE 261

Qy   242 TLLTSIDELA-KAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKA 300
           ||||||||| |:||||||||::   : ||||||:|||||||||||::|||:|
Db   262 TLLASIDELATKAIGKKIQQNGLAVEAGHNGTLLAGAYTISKLITOKLDGLKNSEKLKE 321

Qy   301 EIEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKN 360
           :|| |||||||| | | ||| ||||||||||:|| | |||||||||||||||:|
Db   322 KIENAKKCSEDFTKKLEGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVEN 381

Qy   361 LSKAAKEMLTNSVKEL 376
           |:||||||| |||||
Db   382 LAKAAKEMLANSVKEL 397

Search completed: March 18, 2002, 10:08:43
Job time: 971 sec

```
                    GenCore version 4.5
            Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on:      March 18, 2002, 09:58:16 ; Search time 81.02 seconds
                                         (without alignments)
                                         298.042 Million cell updates/sec

Title:          US-09-596-746a-28
Perfect score:  1859
Sequence:       1 MACNNSGKDGNTSANSADES............KNLSKAAKEMLTNSVKELTS 378

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5
Searched:       303921 seqs, 63882009 residues

Total number of hits satisfying chosen parameters:      303921

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 10%
                 Listing first 45 summaries

Database : Pending_Patents_AA_New:*
    1: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pep:*
    2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep:*
    3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep:*
    4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep:*
    5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep:*
    6: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep1:*
    7: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep:*
    8: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                              SUMMARIES
```

Result No.	Score	Query % Match	Length	DB	ID	Description
1	692.5	37.3	211	6	US-09-974-992-5	Sequence 5, Appli
2	631	33.9	212	6	US-09-974-992-7	Sequence 7, Appli
3	596.5	32.1	207	6	US-09-974-992-3	Sequence 3, Appli
4	171.5	9.2	1881	7	US-10-032-585-7646	Sequence 7646, Ap
5	171.5	9.2	1881	7	US-10-072-851-15590	Sequence 15590, A
6	171	9.2	1313	6	US-09-708-427-15044	Sequence 15044, A
7	170	9.1	1304	6	US-09-708-427-15045	Sequence 15045, A
8	167.5	9.0	1014	6	US-09-708-427-19883	Sequence 19883, A
9	167.5	9.0	1018	6	US-09-708-427-19882	Sequence 19882, A
10	167.5	9.0	1269	6	US-09-708-427-19881	Sequence 19881, A
11	162.5	8.7	837	6	US-09-815-242-5883	Sequence 5883, Ap
12	162.5	8.7	837	6	US-09-708-427-19884	Sequence 19884, A
13	162.5	8.7	875	6	US-09-815-242-13080	Sequence 13080, A
14	162.5	8.7	875	7	US-10-072-851-13080	Sequence 13080, A
15	159	8.7	2434	6	US-09-815-242-5835	Sequence 5835, Ap
16	159	8.6	2434	7	US-10-072-851-5835	Sequence 5835, Ap
17	159	8.6	6281	6	US-09-815-242-12996	Sequence 12996, A
18	159	8.6	6281	7	US-10-072-851-12996	Sequence 12996, A
19	157	8.4	670	5	US-09-748-875-63	Sequence 63, Appl
20	157	8.4	1725	7	US-10-037-182-12	Sequence 12, Appl
21	157	8.4	1786	6	US-09-938-275-7	Sequence 7, Appli
22	157	8.4	1786	7	US-10-037-182-10	Sequence 10, Appl
23	157	8.4	2116	6	US-09-614-150-1122	Sequence 1122, Ap
24	157	8.4	2437	7	US-10-072-851-5834	Sequence 5834, Ap
25	157	8.4	2437	7	US-09-815-242-851-5834	Sequence 5834, Ap
26	156	8.4	1144	6	US-09-708-427-15046	Sequence 15046, A
27	155.5	8.4	589	5	US-09-748-875-14	Sequence 14, Appl
28	155.5	8.4	690	5	US-09-748-875-61	Sequence 61, Appl
29	155.5	8.4	691	5	US-09-748-875-1	Sequence 1, Appli
30	155.5	8.4	701	5	US-09-748-875-62	Sequence 62, Appli
31	155.5	8.4	707	5	US-09-748-875-2	Sequence 2, Appli
32	155.5	8.4	929	5	US-09-748-875-60	Sequence 60, Appl
33	155	8.3	1703	6	US-09-708-427-21161	Sequence 21161, A
34	155	8.3	1710	6	US-09-708-427-21160	Sequence 21160, A
35	155	8.3	1744	6	US-09-708-427-21159	Sequence 21159, A
36	151.5	8.1	2086	6	US-09-815-242-5639	Sequence 5639, Ap
37	151.5	8.1	2086	7	US-10-072-851-5639	Sequence 5639, Ap
38	151.5	8.1	5795	6	US-09-815-242-12610	Sequence 12610, A
39	151.5	8.1	5795	7	US-10-072-851-12610	Sequence 12610, A
40	148	8.0	1690	5	US-09-614-150-10224	Sequence 10224, A
41	148	8.0	1690	5	US-09-614-150-10311	Sequence 10311, A
42	148	8.0	2056	5	US-09-614-150-4824	Sequence 4824, Ap
43	148	8.0	3070	6	US-09-961-403-7	Sequence 7, Appli
44	148	8.0	3110	6	US-09-561-709B-7	Sequence 7, Appli
45	147.5	7.9	1639	6	US-09-614-150-6213	Sequence 6213, Ap

```
                              ALIGNMENTS

RESULT 1
US-09-974-992-5
  Sequence 5, Application US/09974992
; GENERAL INFORMATION:
;  APPLICANT: Mathiesen, Marianne J.
;  APPLICANT: Theisen, Michael
;  APPLICANT: Holm, Arne
;  APPLICANT: Ostergaard, Soren
;  TITLE OF INVENTION: Novel OspC-derived peptide fragments
;  FILE REFERENCE: 459-666P
;  CURRENT APPLICATION NUMBER: US/09/974,992
;  CURRENT FILING DATE: 2001-10-10
;  PRIOR APPLICATION NUMBER: US/09/180,089
;  PRIOR FILING DATE: 1999-05-13
;  PRIOR APPLICATION NUMBER: PCT/DK97/00203
;  PRIOR FILING DATE: 1997-05-02
;  NUMBER OF SEQ ID NOS: 40
;  SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 5
;  LENGTH: 211
;  TYPE: PRT
;  ORGANISM: Borrelia burgdorferi
US-09-974-992-5

Query Match             37.3%;  Score 692.5;  DB 6;  Length 211;
Best Local Similarity   77.0%;  Pred. No. 8.6e-40;
Matches  151;  Conservative  12;  Mismatches  30;  Indels  3;  Gaps  3;
```

```
Qy    1  MACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKA   60
         :||||||||||||||||||||||||||||||| |||||||||||||||||||:||||||:
Db   17  ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEL-AKA   75

Qy   61  IGKKIHIQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLK-NEGLKEKIDAAKKCSETFT  119
         |||||  |  |||||||||| |||||||||||||||| || : |||||||||||||||||:
Db   76  IGKKIKNDGSLGDEANHNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEEFS  135

Qy  120  NKLKEKHTDLGKEGVTDADAKEAILKANGT-KTKGAEELGKLFESVELLSKAAKEMLANS  178
         ||||| || ||| :||||| :|||:||||||   |||: |||||||||||||||||||||
Db  136  TKLKDNHAQLGIQGVTDENARKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANS  195

Qy  179  VKELTSPVVAESPKKP  194
         |||||||||  ||||||
Db  196  VKELTSPVVESPKKP  211

RESULT 2
US-09-974-992-7
```

```
Sequence 7, Application US/09974992
GENERAL INFORMATION:
  APPLICANT: Mathiesen, Marianne J.
  APPLICANT: Theisen, Michael
  APPLICANT: Holm, Arne
  APPLICANT: Ostergaard, Soren
  TITLE OF INVENTION: Novel OspC-derived peptide fragments
  FILE REFERENCE: 459-666P
  CURRENT APPLICATION NUMBER: US/09/974,992
  CURRENT FILING DATE: 2001-10-10
  PRIOR APPLICATION NUMBER: 09/180,089
  PRIOR FILING DATE: 1999-05-13
  PRIOR APPLICATION NUMBER: PCT/DK97/00203
  PRIOR FILING DATE: 1997-05-02
  NUMBER OF SEQ ID NOS: 40
  SOFTWARE: PatentIn Ver. 2.1
  SEQ ID NO 7
  LENGTH: 212
  TYPE: PRT
  ORGANISM: Borrelia afzelii
US-09-974-992-7

Query Match        33.9%;  Score 631;  DB 6;  Length 212;
Best Local Similarity   69.9%;  Pred. No. 1.2e-35;
Matches 137;  Conservative 15;  Mismatches 42;  Indels 2;  Gaps 2;

Qy    1 MACNNSGKDGNT-SANSADESVKGPNLTEINKKITDSNAVLLAVKVEVEALLSSIDEIAAK   59
          ::|||||| |||| | ||| ||:||||||||||||||||||||||||||||::|
Db   17 ISCNNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETTLVLLSIDELAKK   76

Qy   60 AIGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKN-EGLKEKIDAAKKCSEFTN  118
          |||:|| ||| ||| |||| ||||||||||||||||||:| ::| |||||||||||||| |
Db   77 AIGQKIDNNNGLAALNNQNGSLLAGAYAISTLITEKLSKLKNLEELKTEIAKAKKCSEEF   136

Qy  119 TNKLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKRAKEMLANS  178
          |||| ||||| || :|| || :|| ||:|||||||||||||||||||||||||| ||
Db  137 TNKLKSGHADLGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGLLKAAQVALTNS  196

Qy  179 VKELTSPVVAESPKKP  194
          ||||||||||||||||
Db  197 VKELTSPVVAESPKKP  212


RESULT 3
US-09-974-992-3
Sequence 3, Application US/09974992
GENERAL INFORMATION:
  APPLICANT: Mathiesen, Marianne J.
  APPLICANT: Theisen, Michael
  APPLICANT: Holm, Arne
  APPLICANT: Ostergaard, Soren
  TITLE OF INVENTION: Novel OspC-derived peptide fragments
  FILE REFERENCE: 459-666P
  CURRENT APPLICATION NUMBER: US/09/974,992
  CURRENT FILING DATE: 2001-10-10
  PRIOR APPLICATION NUMBER: 09/180,089
  PRIOR FILING DATE: 1999-05-13
  PRIOR APPLICATION NUMBER: PCT/DK97/00203
  PRIOR FILING DATE: 1997-05-02
  NUMBER OF SEQ ID NOS: 40
  SOFTWARE: PatentIn Ver. 2.1
  SEQ ID NO 3
  LENGTH: 207
  TYPE: PRT
  ORGANISM: Borrelia garinii
US-09-974-992-3

Query Match        32.1%;  Score 596.5;  DB 6;  Length 207;
Best Local Similarity   66.0%;  Pred. No. 2.5e-33;
Matches 128;  Conservative 25;  Mismatches 38;  Indels 3;  Gaps 2;
```

```
Qy    1 MACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKVEVEALLSSIDEIAAKA   60
          :|||||| ::| : |||| : |||||||: |||||||||: ||||| ||||||||||: :|
Db   17 ISCNNSG--GDSASTNPDESAKGPNLTVISKKITDSNAFLLAVKEVEALLSSIDEL-SKA   73

Qy   61 IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSEFTN  120
          ||||| ||| ||:|| ::|| ||||| || || ||| | ::| |||||:|| ||| ||
Db   74 IGKKIKNDGTLDNEANRNESLIAGAYEISKLITQKLSVLNSEELKEKIKEAKDCSEKFTT  133

Qy  121 KLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK  180
          |||: ||:||:| : ||:|| ||||:|: ||||||||| ||:||:||||||  |||:|||
Db  134 KLKDSHAELGIQSVQDDNAKKAILKTHGTRDKGAKELELFKSLESLSKAAQAALTNSVK  193

Qy  181 ELTSPVVAESPKKP  194
          |||:||||||||||
Db  194 ELTNPVVAESPKKP  207


RESULT 4
US-10-032-585-7646
Sequence 7646, Application US/10032585
GENERAL INFORMATION:
  APPLICANT: Terry, Roemer D.
  APPLICANT: Bo, Jiang
  APPLICANT: Charles, Boone
  APPLICANT: Howard, Bussey
  TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
  FILE REFERENCE: 10182-005-999
  CURRENT APPLICATION NUMBER: US/10/032,585
  CURRENT FILING DATE: 2001-12-20
  NUMBER OF SEQ ID NOS: 8000
  SOFTWARE: PatentIn version 3.1
  SEQ ID NO 7646
  LENGTH: 1881
  TYPE: PRT
  ORGANISM: Candida albicans
  FEATURE:
  NAME/KEY: MISC_FEATURE
  LOCATION: (1881)..(1881)
  OTHER INFORMATION: X=any amino acid
US-10-032-585-7646

Query Match        9.2%;  Score 171.5;  DB 7;  Length 1881;
Best Local Similarity   21.0%;  Pred. No. 0.0025;
Matches 100;  Conservative 78;  Mismatches 171;  Indels 127;  Gaps 17;

Qy    1 MACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKVEVEALLSSIDEIAAKA   60
          :| |: : : : :|  ::::| |::|| | : ||::::: : : |: : ::|:|: :|| |
Db 1411 LAAQLETKTSNLDSTTMELEKTELELKVRNELTEATSELTKLQDNNQSLTEIEKTKAA  1470

Qy   61 IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNE-----GLKEKIDAAKKC  114
          : | : ::::| :: : :     :: :: | :|:| : |::       |||||      
Db 1471 LTK---SSKDLEVCGNQKSELQDSLKS--VKSELKNFENKYNQETTSLKDEIEKQKE  1523

Qy  115 SEFTTNKLKEKHTDLGKEGVTDADAKEAILKA------------------------NG  148
          | : ::|:  : ::: :|   :: : | :||:                          
Db 1524 IVILQTELKDRISEVEKERAMLSENSETVIKEYSDKIKSLESKINSIKENHSKEITTPHNE  1583

Qy  149 TKTRGAEELGKLFESVEVLSKAAKEMLANSVKELTSPVVAESPKKPSMVNNSGKDGNTSA  208
            : : | :|| :::  : |: : : ::| : :::                         
Db 1584 QKTSLKQDIAKLSQDHE-SAQTQLEDKENQLKEL------KASL----EKHNTESA  1628

Qy  209 NSADESVKGPNLTEISKKI-------------------------TESNAVVLAVKEVETLLTSIDELAKAIGK  256
          :| ::: : ::: :|:::                                        
Db 1629 TSIEE--KNNQIKELSEFIKSLKTELKTSGDALKQSOKEYKTLKTKNSDTESKLEKQLEE  1686

Qy  257 --KIKNDVSLDNEADIHNGSLISGAYLISNLITKKISAIKDSGELKA-------EIEKAK--  306
           :: ::::|  : | ::| | :: : ::| :   ::  :: || |         :|:|||  
Db 1687 LEKVKSDLQTADE-KLKGITEREIALKSELETVKNSGLSTTSELAALTKTVKSLEKEKEE  1745

Qy  307 -------------KCSEEFTAKLKGEHTDLGK--EGVTD-------------DNAKKAILKTNNDKTK  346
                        ||::|:  :||| ::|||           ::| : ||: |:|:
Db 1746 LQFLSGNKSKELEDYIQK----HSDISEKLKALTDELKEKTKQFDDSKKILTELENDLTS  1801
```

```
Qy   347 GADELE------------------------------------------KLFESVKNLSKAAKEMLTNSVKELTS 378
           |||                                          |  || |  | ||| | |  | ||  |
Db  1802 TKKELETEKTQTSKFKNLEERKDKEIVKLNKELELLKNDNSGAKKELSEKVSKLES 1857
```

RESULT 5
US-10-072-851-15590
Sequence 15590, Application US/10072851
GENERAL INFORMATION:
; APPLICANT: Carr, Grant J.
; APPLICANT: Xu, H. Howard
; APPLICANT: Foulkes, J Gordon
; APPLICANT: Zamudio, Carlos
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Roemer, Terry
; APPLICANT: Jiang, Bo
; APPLICANT: Bussey, Howard
; APPLICANT: Boone, Charles
; ORGANISM: Candida albicans
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (1881)..(1881)
; OTHER INFORMATION: X=any amino acid
US-10-072-851-15590

; TITLE OF INVENTION: Methods for Identifying the Target of a Compound which Inhibits C
; TITLE OF INVENTION: Proliferation
; FILE REFERENCE: ELITRA.028A
; CURRENT APPLICATION NUMBER: US/10/072,851
; CURRENT FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 15811
; SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 15590
; LENGTH: 1881
; TYPE: PRT

Query Match   9.2%; Score 171.5; DB 7; Length 1881;
Best Local Similarity 21.0%; Pred. No. 0.0025;
Matches 100; Conservative 78; Mismatches 171; Indels 127; Gaps 17;

```
Qy     1 MACNNSGCKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKA 60
Db  1411 LAAQLETKTSNLDSTTMELEKTELELKKVRNELTEATSELTKLQODNNQSLTEEIEKTKAA 1470

Qy    61 IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNE------GLEKIDAARKKC 114
Db  1471 LTK---SSKDLEVCGNQKSELCQDSLKS-----VKSELKNFENKYNQEFTSLKDEIEEKOKE 1523

Qy   115 SETFTKLKEKHTDLGKEGVTDADAKEAILKA----------------NG 148
Db  1524 IVTLQTELKDRISEVEKERAMLSENSETVIKEYSDKIKSLESKINSIKENHSKEITTHNE 1583

Qy   149 TKTKGABELGKLFESVEVLSKRAAKEMLANSVKELTSPVVAESPKKPSMVNNSGKDGNTSA 208
Db  1584 OKTSLKQDIAKLSQDHE-SAQTQLEDKENQLKEL------KASL----EKHNTESA 1628

Qy   209 NSADESVKGPNLTEISKKI----------TESNAVVLAVKEVETLLTSIDELAKAIGK---- 256
Db  1629 TSIEE--KNNQIKELSETIKSLKTELKTSGDALKQSOKEYKYLKTKNSDTESKLEKQLEE 1686

Qy   257 --KIKNDVSLDNEADINGSLISGAYLISNLITKKISAIKDSGELKA------EIEKAK-- 306
Db  1687 LEKVKSDLQTADE-KLKGITEREIALKSELETVKNSGLSTTSELAALTKTVKSLEKEKEE 1745

Qy   307 ------------------------------------KCSEEFTAKLLKGEHTDLGK--EGVTD----------DNAKKAILKTNNDKTK 346
```

Db 1746 LQFLSGNKSKELEDYIQK------------HSDISEKLKALTDELKEKTKQPFDDSKKLITELENDLTS 1801

```
Qy   347 GADELE------------------------------------------KLFESVKNLSKAAKEMLTNSVKELTS 378
           |||                                          |  || |  | ||| | |  | ||  |
Db  1802 TKKELETEKTQTSKFKNLEERKDKEIVKLNKELELLKNDNSGAKKELSEKVSKLES 1857
```

RESULT 6
US-09-708-427-15044
Sequence 15044, Application US/09708427
GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
; TITLE OF INVENTION: THEREBY
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: Patentin version 3.1
SEQ ID NO 15044
; LENGTH: 1313
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..1313
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..1313
; OTHER INFORMATION: Ceres Seq. ID 1828627
US-09-708-427-15044

Query Match   9.2%; Score 171; DB 6; Length 1313;
Best Local Similarity 21.8%; Pred. No. 0.0016;
Matches 115; Conservative 61; Mismatches 148; Indels 204; Gaps 21;

```
Qy     8 KDGNTSANSADESVKGPNLTEINKKITDSNAV------LLAVKEVEALLSSIDEIAAKA 60
Db     6 KTGLMETPRSKPSPEPPRLSKLSASKSDSNSASPKVHSRLVKGTELQTLNQIQEDLKKA 65

Qy    61 IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQK--LDGLK-------NEGLKEKI 108
Db   102 AAQKRAEESPEVEKFRAVELEQAGLEAVQKKDVTSKNELESIRSQHALDISALLSTTEEL 161

Qy   109 DAARKCSEFF---------------------------TNK-----LKEKHT-DLG-- 130
Db   162 QRVKHELSMTADARKNKALSHAEEATKIAEIHAEKAEILASELGRLKALLGSKEKREAIEG 221

Qy   131 ----KEGVTDADAKEAILKANGTKTKGAE---------ELGKL----------FES 163
Db   222 NEIVSKLKSEIELLRGELEKVSILESSLKEQEGLVEOLKVDLEAAKMAESCTNSSVEEWK 281

Qy   164 VEVLSKAAKE----------MLANSVKE------------LTSPVVAESP 191
Db   282 NKVHELEKEVEESNRSKSSASESME-----SVMKQLAELNHVLHETKSDNAAQKEIELL 336

Qy   192 KKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVV----------LA 236
Db   337 EKTIEAQRDLEEYGRQVCIAKEEASKLENLVESIKSELEISQEEKTRALDNEKAATSN- 395

Qy   237 ----VKEVETLLTSIDE---------LAKAIGKKIKNDV-----SLDNEADHNGSL 274
Db   396 ISGAYLISNLITKKISAIKDSGELKAEIEKAKCCSEEFTAKLKGEHTDLGKEGVTDNAK 441

Qy   275 ISGAYLISNLITKKISAIKDSGELKAEIEKAKCCSEEFTAKLKGEHTDLGKEGVTDNAK------QEASTESSEA 335
Db   442 KATLLVCQEELKNC--ESQVDSLKLASKETNEKVEKMLEDARNEIDS 486
```

RESULT 7
US-09-708-427-15045
Sequence 15045, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: THEREBY
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; SOFTWARE: Patentin version 3.1
; NUMBER OF SEQ ID NOS: 85364
; SEQ ID NO 15045
; LENGTH: 1304
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..1304
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..1304
; OTHER INFORMATION: Ceres Seq. ID 1828628
US-09-708-427-15045

Query Match   9.1%; Score 170; DB 6; Length 1304;
Best Local Similarity   21.9%; Pred. No. 0.0019;
Matches 112; Conservative 60; Mismatches 136; Indels 204; Gaps 21;

```
Qy  24 PNLTELNKKITDSNAV--------LLAVKEVEALLSSIDEIAAKAIGKKIHQNNGLDTENN 76
Db  13 PRLSKLSASKSDSNSASPKVHSRLVKGTELQTQLNIQEDLKKA----------DEQ-- 59

Qy  77 HNGSLLAGAYAISTLIKQK----LDGLK------NEGLKEKIDAAKKCSETF------- 118
Db  60 ----IELLKKDKAKAIDDLKEUSEKLVEEANEKLKEALAAQKRAEESFEVEKFR 108

Qy 119 --------------TNK----LKEKHT-DLG----------KEGVTDADAKEA 142
Db 109 AVELEQAGLEAVQKKDVTSKNELESIRSQHALDISALLSTTEELQRVKHELSMTADAKNK 168

Qy 143 ILKANGTKTKGAE----------ELGKL------FESVEVLSKAAKE----- 173
Db 169 ALSHAEEATKIAEIHAEKAEIIASELGRIKALLGSKEEKEAIEGNEIVSKLKSEIELLRG 228

Qy 174 ------MLANSYKE----------LTSPVVAESPKKPSMVNNSGKDGNTS 207
Db 229 ELEKVSILESSLKEQEGLVEQLKDVDLEAAKMAESCTNSSVEEWKNKVHELEKEVEESNRS 288

Qy 208 ANSADESVKGPNLTEISKKITESNAVV-----------LAVKEVEFLLTSIDE--- 249
Db 289 KSSASESME----SVNKQLAELNHVLHEFKSDNAAQKEKIELLEKTIEAQRTDLEEYGR 343

Qy 250 ---LAKAIGKKIKNDV------------SLDNEADHNGSLISGAYLISNLITKKIS 290
Db 344 OVCIAKEEASKLENLVESIKSELEISQQEEKTRALDNEKAAISN--- 396

Qy 291 AIKDSGELKAEIEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADE 350
Db 397 LSIELERCKVEEEKSKKDMESLILAL------QEASTESSEAKATLLVCQEELKNC-- 446

Qy 351 LEKLFESVKNLSKAAKE----MLINSVKELTS 378
Db 447 -ESQVDSLIKLASKETNEKYEKMLEDARNEIDS 477
```

RESULT 8
US-09-708-427-19883
Sequence 19883, Application US/09708427
; GENERAL INFORMATION:

; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
; TITLE OF INVENTION: THEREBY
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; SOFTWARE: Patentin version 3.1
; NUMBER OF SEQ ID NOS: 85364
; SEQ ID NO 19883
; LENGTH: 1014
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..1014
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..1014
; OTHER INFORMATION: Ceres Seq. ID 1836847
US-09-708-427-19883

Query Match   9.08%; Score 167.5; DB 6; Length 1014;
Best Local Similarity   23.18%; Pred. No. 0.002;
Matches 113; Conservative 79; Mismatches 155; Indels 143; Gaps 23;

```
Qy   8 KDGNTSANSADE----------------SVKGPNL-------TEINKKITDS-NAVLLAVKE 45
Db 401 KHGETEADSKGYLGQVAELQSTLEAFQVKSSSLEAALNIATENEKELIENLNAVTSEKKK 460

Qy  46 VEALLSSIDEIAAK----------AIGKKIHQNNGLDTENNHNGSLLAGAYAISTL--IKQ 94
Db 461 LEA----TVDEYSVKISESENLLESTRNELNVTQG-KLESIENDLKAAGLQESEVMEKLKS 516

Qy  95 KIDGLKNEGLKEKIDAA------------------KKCSETFTNK-----LKEKH 126
Db 517 AEESLEQKG--REIDEATTKRMELEALHQSLSIDSEHRLQKAMEEFTSRDSEASSLTEKL 574

Qy 127 TDL-GKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKEL--- 182
Db 575 RDLEGKISYEEQLAEASGKSSSLKEKLEQTLGRLAAAESVNEKLKQEFDQAQEKSLQSS 634

Qy 183 -TSPVVAESPKK--------PSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAV 233
Db 635 SESELLAETNNQLKIKIQELEGELIGSGSVEKEFALKRLEEALE----RFNQKETESSDL 689

Qy 234 VLAVKEVEFLLTSIDELA-KAIG---------------KKIKN--- 260
Db 690 VEKLKTHENQIEEYKRLAHEASGVADTRKVELEDALSKLKNLESTIEELGAKCQGLEKES 749

Qy 261 -----DVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKA-EIKKAKKCSEEFT 313
Db 750 GDIAEVNLKLNELANIGS----EANELQTKLSALEAEKEQTANELEASKTTIEDLT 802

Qy 314 AKL--KGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVK-NLSKAAKEMLT 370
Db 803 KQLTSEGEKLQSQIEKLRAVAAEKSVLESH-----FEELEKTLSEVKAQLKENVENAAT 856

Qy 371 NSVK---ELTS 378
Db 857 ASVKVAELTS 866
```

RESULT 9
US-09-708-427-19882
Sequence 19882, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
; TITLE OF INVENTION: THEREBY
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09

NUMBER OF SEQ ID NOS: 85364
SOFTWARE: Patentin version 3.1
SEQ ID NO 19882
LENGTH: 1018
TYPE: PRT
ORGANISM: Arabidopsis thaliana
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1..1018
OTHER INFORMATION:
NAME/KEY: misc_feature
LOCATION: 1..1018
OTHER INFORMATION: Xaa is any amino acid
LOCATION: 1..1018
OTHER INFORMATION: Ceres Seq. ID 1836846
US-09-708-427-19882

Query Match 9.0%; Score 167.5; DB 6; Length 1018;
Best Local Similarity 23.1%; Pred. No. 0.002;
Matches 113; Conservative 79; Mismatches 155; Indels 143; Gaps 23;

```
Qy    8 KDGNTSANSADE------------------------SVKGPNL------------TEINKKITDS-NAVLLAVKE 45
          |:|                              :|              ||| :|:: ||| :
Db  405 KHGETEADSKGYLGQVAELQSTLEAFQVKSSSLEAALNIATENEKELTENLNAVTSEKKK 464

Qy   46 VEALLSSIDEIAAK-------------------AIGKKIHQNNGLDTENNHNGSLLAGAYAISTL--IKQ 94
          | :|                          |:| ||
Db  465 LEA---TVDEYSVKISESENLLESIRNELNVTQG-KLESIENDLKAAGLQESEVMEKLKS 520

Qy   95 KLDGLKNEGLKEKIDAA----------------KKCSETFTNK-------LKEKH 126
          |||   |||:|||                  |  :| |||        ||||
Db  521 AEESLEQKG--REIDEATTKMELEALHQSLSIDSEHRLQKAMEEFTSRDSEASSLTEKL 578

Qy  127 TDL-GKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKEL--- 182
          |||  |||  |:| |||  |||| |||
Db  579 RDLEGKIKSYEEQLAEASGKSSSLKEKLEQTLGRLAAAESVNEKLKQEFDQAQEKSLQSS 638

Qy  183 -TSPVVAESPKK-------------PSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAV 233
           :||    :|             ||| |||  ||
Db  639 SESELLAETNNQLKIKIQELEGLIGSGSVEKETALKRLEEAIE------RFNQKETESSDL 693

Qy  234 VLAVKEVETLLTSIDELA-KAIG------------KKIKN------ 260
          ||||||| ||||||||| ||||            |||||
Db  694 VEKLKTHENQIEEYKKLAHEASGVADTRKVELEDALSKLKNLESTIEELGAKCQGLEKES 753

Qy  261 -----DVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKA-EIEKAKKCSEEFT 313
               |||| |||||| |||  ||| :: |||      |:||||
Db  754 GDLAEVNLKLNLELANHGS------EANELQTKLSALEAEKEQTANELEASKTTIEDLT 806

Qy  314 AKL--KGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVK-NLSKAAKEMLT 370
          |||  ||| ||| |||:||||           |:| ||||  ||||| :||||||||:|
Db  807 KQLTSEGEKLQSQIEKLRAVAAEKSVLESH-----FEELEKTLSEVKAQLKENVENAAT 860

Qy  371 NSVK--ELTS 378
          |:||  ||:|
Db  861 ASVKVAELTS 870
```

RESULT 10
US-09-708-427-19881
Sequence 19881, Application US/09708427
GENERAL INFORMATION:
APPLICANT: N. ALEXANDROV et al.
TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
FILE REFERENCE: 2750-1243P
CURRENT APPLICATION NUMBER: US/09/708,427
CURRENT FILING DATE: 2000-11-09
NUMBER OF SEQ ID NOS: 85364
SOFTWARE: Patentin version 3.1
SEQ ID NO 19881
LENGTH: 1269
TYPE: PRT
ORGANISM: Arabidopsis thaliana

FEATURE:
NAME/KEY: misc_feature
LOCATION: 1..1269
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: 1..1269
OTHER INFORMATION: Ceres Seq. ID 1836845
US-09-708-427-19881

Query Match 9.0%; Score 167.5; DB 6; Length 1269;
Best Local Similarity 23.1%; Pred. No. 0.0027;
Matches 113; Conservative 79; Mismatches 155; Indels 143; Gaps 23;

```
Qy    8 KDGNTSANSADE------------------------SVKGPNL------------TEINKKITDS-NAVLLAVKE 45
          |:|                              :|              ||| :|:: ||| :
Db  656 KHGETEADSKGYLGQVAELQSTLEAFQVKSSSLEAALNIATENEKELTENLNAVTSEKKK 715

Qy   46 VEALLSSIDEIAAK-------------------AIGKKIHQNNGLDTENNHNGSLLAGAYAISTL--IKQ 94
          | :|                          |:| ||
Db  716 LEA---TVDEYSVKISESENLLESIRNELNVTQG-KLESIENDLKAAGLQESEVMEKLKS 771

Qy   95 KLDGLKNEGLKEKIDAA----------------KKCSETFTNK-------LKEKH 126
          |||   |||:|||                  |  :| |||        ||||
Db  772 AEESLEQKG--REIDEATTKMELEALHQSLSIDSEHRLQKAMEEFTSRDSEASSLTEKL 829

Qy  127 TDL-GKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKEL--- 182
          |||  |||  |:| |||  |||| |||
Db  830 RDLEGKIKSYEEQLAEASGKSSSLKEKLEQTLGRLAAAESVNEKLKQEFDQAQEKSLQSS 889

Qy  183 -TSPVVAESPKK-------------PSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAV 233
           :||    :|             ||| |||  ||
Db  890 SESELLAETNNQLKIKIQELEGLIGSGSVEKETALKRLEEAIE------RFNQKETESSDL 944

Qy  234 VLAVKEVETLLTSIDELA-KAIG------------KKIKN------ 260
          ||||||| ||||||||| ||||            |||||
Db  945 VEKLKTHENQIEEYKKLAHEASGVADTRKVELEDALSKLKNLESTIEELGAKCQGLEKES 1004

Qy  261 -----DVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKA-EIEKAKKCSEEFT 313
               |||| |||||| |||  ||| :: |||      |:||||
Db 1005 GDLAEVNLKLNLELANHGS------EANELQTKLSALEAEKEQTANELEASKTTIEDLT 1057

Qy  314 AKL--KGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVK-NLSKAAKEMLT 370
          |||  ||| ||| |||:||||           |:| ||||  ||||| :||||||||:|
Db 1058 KQLTSEGEKLQSQIEKLRAVAAEKSVLESH-----FEELEKTLSEVKAQLKENVENAAT 1111

Qy  371 NSVK--ELTS 378
          |:||  ||:|
Db 1112 ASVKVAELTS 1121
```

RESULT 11
US-09-815-242-5883
Sequence 5883, Application US/09815242
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
Prokaryotes
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727

RESULT 12
US-10-072-851-5883
Sequence 5883, Application US/10072851
GENERAL INFORMATION:
APPLICANT: Carr, Grant J.
APPLICANT: Xu, H. Howard
APPLICANT: Foulkes, J Gordon
APPLICANT: Zamudio, Carlos
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Roemer, Terry
APPLICANT: Jiang, Bo
APPLICANT: Boone, Charles
APPLICANT: Bussey, Howard
TITLE OF INVENTION: Methods for Identifying the Target of a Compound which Inhibits C
TITLE OF INVENTION: Proliferation
FILE REFERENCE: ELITRA.028A
CURRENT APPLICATION NUMBER: US/10/072,851

; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSEQ for Windows Version 4.0

SEQ ID NO 5883
; LENGTH: 837
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-5883

Query Match   8.7%; Score 162.5; DB 6; Length 837;
Best Local Similarity 18.7%; Pred. No. 0.0033;
Matches 76; Conservative 74; Mismatches 181; Indels 75; Gaps 12;

Qy   9 DGNTSANSADESVKGP--NLTEINK----------------KITDSNAVLLAVKEVEALLSS 52
            ||| || ::| | ||| || | | |:   ::
Db   350 NGNQNVANAKTTAKNALNNLTSINNAQKEALKSOIEGATTVAGVNQVSTTASELNTAMSN 409

Qy   53 I-----DEIAAKAIGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEK 107
                  |:| || | | | | ::|| || | | | ||:| | :
Db   410 LQNGINDEAATKA---ALNGTQNLEKAKQHANTAIDGL------SHLTNAQKEALKQL 458

Qy   108 IDAAKKCSETFTNKLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVL 167
            :| ::: | | |: | ||||| ||| || | || | | ||| |||| ||
Db   459 VQQSTTVAEAQGNEQKANNVDAAMDKLRQSIADNATTKQNQNYTDASQNKKDAYNNAVTT 518

Qy   168 SKAAKEMLANSVKELTSPVVAESPKKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKI 227
            | | | | | | | | | : | | | |:| || || | | | ||| |
Db   519 AQGI-----IDQTISPTL-----DPTIVINQAAGQVSTTKNALNGN--ENLEAAKQQA 563

Qy   228 TESNAVVLAVKEVETLLTSIDELAKAIGKKIKNDVSLDNEADHNGSLLISGAYLISNLITK 287
            | | |:   :: | |:|| || | || | | | :| | | || | | |:| |
Db   564 SQS----------LGSLDNLNNAQKQTVTDQINGAHTVDEANQIKQNAQNLNTAMGN 610

Qy   288 KISAIKDSGELKAEI---------EKAKK--------CSEEFTAKILKGEHTDLG--KEGVTDDN 332
            | | |:|| |: : ||| | ||: | | | | || |||| |
Db   611 LKQAIADKDATKATVNFTDADQAKQQAYNTAVTNAENIISKANGGNATQAEVEQAIKQVN 670

Qy   333 AKKAILKTNNDKTKGADELEKLFESVKNLSKAAKEMLTNSVKELTS 378
            ||:|| |||:| |||||||||||||| ||||||| | ||||||||
Db   671 AAKQALNGNANVQHAKDEATALINSSNDLNQAQKDALKQQVQNATT 716

RESULT 13
US-09-815-242-13080
Sequence 13080, Application US/09815242
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
TITLE OF INVENTION: Prokaryotes
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; CURRENT APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931

; CURRENT FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 15811
; SOFTWARE: FastSEQ for Windows Version 4.0

SEQ ID NO 5883
; LENGTH: 837
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-072-851-5883

Query Match   8.7%; Score 162.5; DB 7; Length 837;
Best Local Similarity 18.7%; Pred. No. 0.0033;
Matches 76; Conservative 74; Mismatches 181; Indels 75; Gaps 12;

Qy   9 DGNTSANSADESVKGP--NLTEINK----------------KITDSNAVLLAVKEVEALLSS 52
            ||| || ::| | ||| || | | |:   ::
Db   350 NGNQNVANAKTTAKNALNNLTSINNAQKEALKSOIEGATTVAGVNQVSTTASELNTAMSN 409

Qy   53 I-----DEIAAKAIGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEK 107
                  |:| || | | | | ::|| || | | | ||:| | :
Db   410 LQNGINDEAATKA---ALNGTQNLEKAKQHANTAIDGL------SHLTNAQKEALKQL 458

Qy   108 IDAAKKCSETFTNKLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVL 167
            :| ::: | | |: | ||||| ||| || | || | | ||| |||| ||
Db   459 VQQSTTVAEAQGNEQKANNVDAAMDKLRQSIADNATTKQNQNYTDASQNKKDAYNNAVTT 518

Qy   168 SKAAKEMLANSVKELTSPVVAESPKKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKI 227
            | | | | | | | | | : | | | |:| || || | | | ||| |
Db   519 AQGI-----IDQTISPTL-----DPTIVINQAAGQVSTTKNALNGN--ENLEAAKQQA 563

Qy   228 TESNAVVLAVKEVETLLTSIDELAKAIGKKIKNDVSLDNEADHNGSLLISGAYLISNLITK 287
            | | |:   :: | |:|| || | || | | | :| | | || | | |:| |
Db   564 SQS----------LGSLDNLNNAQKQTVTDQINGAHTVDEANQIKQNAQNLNTAMGN 610

Qy   288 KISAIKDSGELKAEI---------EKAKK--------CSEEFTAKILKGEHTDLG--KEGVTDDN 332
            | | |:|| |: : ||| | ||: | | | | || |||| |
Db   611 LKQAIADKDATKATVNFTDADQAKQQAYNTAVTNAENIISKANGGNATQAEVEQAIKQVN 670

Qy   333 AKKAILKTNNDKTKGADELEKLFESVKNLSKAAKEMLTNSVKELTS 378
            ||:|| |||:| |||||||||||||| ||||||| | ||||||||
Db   671 AAKQALNGNANVQHAKDEATALINSSNDLNQAQKDALKQQVQNATT 716

;   PRIOR FILING DATE: 2000-12-22
;   PRIOR APPLICATION NUMBER: 60/269,308
;   PRIOR FILING DATE: 2001-02-16
;   NUMBER OF SEQ ID NOS: 14110
;   SOFTWARE: FastSEQ for Windows Version 4.0
;   SEQ ID NO 13080
;   LENGTH: 875
;   TYPE: PRT
;   ORGANISM: Staphylococcus aureus
US-09-815-242-13080

Query Match   8.7%;   Score 162.5;   DB 6;   Length 875;
Best Local Similarity 18.7%;   Pred. No. 0.0035;
Matches 76; Conservative 74; Mismatches 181; Indels 75; Gaps 12;

Qy   9 DGNTSANSADESVKGP--NLTEINK----------------------KITDSNAVLLAVKEVEALLSS 52
         :||: : :|: |||:|| :| |:|:||||||| |::|
Db   387 NGNQNVANAKTTAKNALNNLTSINNAQKEALKSQIEGATTVAGVNQVSTTASELNTAMSN 446

Qy   53 I-----DEIAAKAIGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEK 107
         | |:| || | |:| |||:|:|||| |:|
Db   447 LONGINDEAATKA---ALNGTQNLEKAKQHANTAIDGL-------SHLTNAQKEALKQL 495

Qy   108 IDAAKKCSETFTNKLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVL 167
         : :| | :: |: |||:||
Db   496 VQQSTTVAEAQGNEQKANNVDAAMDKLRQSIADNATTKQNQNYDASQNKKDAYNNAVTT 555

Qy   168 SKAAKEMLANSVKELTSPVVAESPKKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKI 227
         |||||
Db   556 AQGI------IDQTTSPTL----DPTVINQAAGQVSTTKNALNGN--ENLEAAKQQA 600

Qy   228 TESNAVVLAVKEVETLLTSIDELAKAIGKKIKNDVSLDNEADHGNSLISGAYLISNLITK 287
         ::|
Db   601 SQS----------LGSLDNLNNAQKQTVTDQINGAHTVDEANQIKQNAQNLNTAMGN 647

Qy   288 KISAIKDSGELKAEI-------EKAKK-------CSEEFTAKLKGEHTDLG--KEGVTDDN 332
         |:| :: :|
Db   648 LKQAIADKDATKATVNFTDADQAKQQAYNTAVTNAENIISKANGGNATQAEVEQAIKQVN 707

Qy   333 AKKAILKTNNDKTKGGADELEKLFESVKNLSKAAKEMLTNSVKELTS 378
         | || |: ||::
Db   708 AAKQALNGNANVQHAKDEATALINSSNDLNQAQKDALKQQVQNATT 753

RESULT 14
US-10-072-851-13080
; Sequence 13080, Application US/10072851
; GENERAL INFORMATION:

;   APPLICANT: Carr, Grant J.
;   APPLICANT: Foulkes, J. Gordon
;   APPLICANT: Zamudio, Carlos
;   APPLICANT: Haselbeck, Robert
;   APPLICANT: Ohlsen, Karl L.
;   APPLICANT: Zyskind, Judith W.
;   APPLICANT: Wall, Daniel
;   APPLICANT: Trawick, John D.
;   APPLICANT: Yamamoto, Robert T.
;   APPLICANT: Roemer, Terry
;   APPLICANT: Jiang, Bo
;   APPLICANT: Boone, Charles
;   APPLICANT: Bussey, Howard
;   TITLE OF INVENTION: Methods for Identifying the Target of a Compound which Inhibits C
;   TITLE OF INVENTION: Proliferation
;   FILE REFERENCE: ELITRA.028A
;   CURRENT APPLICATION NUMBER: US/10,072,851
;   CURRENT FILING DATE: 2002-02-08
;   PRIOR APPLICATION NUMBER: 60/267,636
;   PRIOR FILING DATE: 2001-02-09
;   NUMBER OF SEQ ID NOS: 15811
;   SOFTWARE: FastSEQ for Windows Version 4.0
;   SEQ ID NO 13080
;   LENGTH: 875
;   TYPE: PRT
;   ORGANISM: Staphylococcus aureus
US-10-072-851-13080

Query Match   8.7%;   Score 162.5;   DB 7;   Length 875;
Best Local Similarity 18.7%;   Pred. No. 0.0035;
Matches 76; Conservative 74; Mismatches 181; Indels 75; Gaps 12;

Qy   9 DGNTSANSADESVKGP--NLTEINK----------------------KITDSNAVLLAVKEVEALLSS 52
         :||: : :|: |||:|| :| |:|:||||||| |::|
Db   387 NGNQNVANAKTTAKNALNNLTSINNAQKEALKSQIEGATTVAGVNQVSTTASELNTAMSN 446

Qy   53 I-----DEIAAKAIGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEK 107
         | |:| || | |:| |||:|:|||| |:|
Db   447 LONGINDEAATKA---ALNGTQNLEKAKQHANTAIDGL-------SHLTNAQKEALKQL 495

Qy   108 IDAAKKCSETFTNKLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVL 167
         : :| | :: |: |||:||
Db   496 VQQSTTVAEAQGNEQKANNVDAAMDKLRQSIADNATTKQNQNYDASQNKKDAYNNAVTT 555

Qy   168 SKAAKEMLANSVKELTSPVVAESPKKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKI 227
         |||||
Db   556 AQGI------IDQTTSPTL----DPTVINQAAGQVSTTKNALNGN--ENLEAAKQQA 600

Qy   228 TESNAVVLAVKEVETLLTSIDELAKAIGKKIKNDVSLDNEADHGNSLISGAYLISNLITK 287
         ::|
Db   601 SQS----------LGSLDNLNNAQKQTVTDQINGAHTVDEANQIKQNAQNLNTAMGN 647

Qy   288 KISAIKDSGELKAEI-------EKAKK-------CSEEFTAKLKGEHTDLG--KEGVTDDN 332
         |:| :: :|
Db   648 LKQAIADKDATKATVNFTDADQAKQQAYNTAVTNAENIISKANGGNATQAEVEQAIKQVN 707

Qy   333 AKKAILKTNNDKTKGGADELEKLFESVKNLSKAAKEMLTNSVKELTS 378
         | || |: ||::
Db   708 AAKQALNGNANVQHAKDEATALINSSNDLNQAQKDALKQQVQNATT 753

RESULT 15
US-09-815-242-5835
; Sequence 5835, Application US/09815242
; GENERAL INFORMATION:

;   APPLICANT: Haselbeck, Robert
;   APPLICANT: Ohlsen, Karl L.
;   APPLICANT: Zyskind, Judith W.
;   APPLICANT: Wall, Daniel
;   APPLICANT: Trawick, John D.
;   APPLICANT: Carr, Grant J.
;   APPLICANT: Yamamoto, Robert T.
;   APPLICANT: Xu, H. Howard
;   TITLE OF INVENTION: Identification of Essential Genes in
;   TITLE OF INVENTION: Prokaryotes
;   FILE REFERENCE: ELITRA.011A
;   CURRENT APPLICATION NUMBER: US/09,815,242
;   CURRENT FILING DATE: 2001-03-21
;   PRIOR APPLICATION NUMBER: 60/191,078
;   PRIOR FILING DATE: 2000-03-21
;   PRIOR APPLICATION NUMBER: 60/206,848
;   PRIOR FILING DATE: 2000-05-23
;   PRIOR APPLICATION NUMBER: 60/207,727
;   PRIOR FILING DATE: 2000-05-26
;   PRIOR APPLICATION NUMBER: 60/242,578
;   PRIOR FILING DATE: 2000-10-23
;   PRIOR APPLICATION NUMBER: 60/253,625
;   PRIOR FILING DATE: 2000-11-27
;   PRIOR APPLICATION NUMBER: 60/257,931
;   PRIOR FILING DATE: 2000-12-22
;   PRIOR APPLICATION NUMBER: 60/269,308
;   PRIOR FILING DATE: 2001-02-16
;   NUMBER OF SEQ ID NOS: 14110
;   SOFTWARE: FastSEQ for Windows Version 4.0
;   SEQ ID NO 5835

```
;   LENGTH: 2434
;   TYPE: PRT
;   ORGANISM: Staphylococcus aureus
US-09-815-242-5835

Query Match     8.6%;   Score 159;   DB 6;   Length 2434;
Best Local Similarity   19.7%;   Pred. No. 0.025;
Matches 105;   Conservative 75;   Mismatches 182;   Indels 172;   Gaps 18;

Qy      4 NNSGK-DGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEI------ 56
            ||:|:||  :|               ::|: :|:|  :||:   || :| ::   ::||     :|
Db   1856 NNAQKQDLNTQVTNA----------PNISTVNQVTKAEQLDQAMERLINGIQDKDQVKQSVNF 1909

Qy     57 ----AAKAIGKKIHQNNGLDTENNHNGSLLAGAYAISTLIIKQKLDGLK--- 100
             |   ||| :    : |  ||| :| :    |    :    | :|:|| : :|:
Db   1910 TDADPEKQTAYNNAVTRAENIINQANG---TNANQSQVEAALSTVTTTKQALNGDRKVT 1965

Qy    101 ------NEGLK--EKIDAAAKCSET----------------------FTNKLKE 124
            |:|  :|:  |:  | : ||               |:|:|:
Db   1966 DAKNNANQTLSTLDNLNNAQKGAVTGNINQAHTVAEVTQAIQTAQELNTAMGNLKNSLND 2025

Qy    125 KHTDLGKEGVTDADAK--------EAILKA---------N 147
            :|  | :| | : |  :         | ::|:  |
Db   2026 KDTTLGSQNFADADPEKKNAYNEAVHNAENILNKSTGTNVPKDQVEAAMNQVNATKAALN 2085

Qy    148 GTKT--KGAEELGKLFESVEVLSKAAKEMLANSVKELTSPVVAESPKKPSMVNN----SG 201
            |:|:  ::||:::| :  ::|  :| : |:|       :|: | |:
Db   2086 GTQNLEKAKQHANTAIDGLSHLTNAQKEALKQLVQQSTT--VAEAQGNEQKANNVDAAMD 2143

Qy    202 KDGNTSANSADESVKGPNLTEIS--KRITESNAVV----------- 234
            |:|:| ||  : |:  |:   ::|   |:|  |
Db   2144 KLRQSIADNA-TTKQNQNYTDASQNKDAYNNAVTTRAQGIIDQTTSPTLDPTVINQAAGQ 2202

Qy    235 ------LAVKEVETLLTSIDELAKAIGKKIKNDVSLDNEADHNGSLISGAY 279
             :| ::|: |:   | :| : ||:|||:|:|:|
Db   2203 VSTTKNALNGNENLEAAKQQASQSLGSLDNLNNAQKQTVTDQINGAHTVDEANQIKQNAQ 2262

Qy    280 LISNLITKKISAIKDSGELKAEI----EKAKK------CSEEFTAKLKGEHTDLG-- 324
             ::|  ||| :  :|    ::|::   |:|||                | ::|: |
Db   2263 NLNTAMGNLKQAIADKDATKATVNFTDADQAKQQAYNTAVTNAENIISKANGGNATQAEV 2322

Qy    325 KEGVTDDNAKRAILKTNNDKTKGADELEKLFPSVKNLSKAAKEMLTNSVKELTS 378
             :|:  | |:| | |   |  | :::|:|::|:|||:||:|  ::: :::
Db   2323 EQAIKQVNAAKQALNGNANVQHAKDEATALINSSNDLNQAQKDALKQQVQNATT 2376
```

OM protein - protein search, using sw model

Copyright (c) 1993 - 2000 Compugen Ltd.   GenCore version 4.5

Run on: March 18, 2002, 10:11:42 ; Search time 39.62 Seconds (without alignments) 349.806 Million cell updates/sec

Title: US-09-596-746A-28
Perfect score: 1859
Sequence: 1 MACNNSGKDGNTSANSADES.............KNLSKAAKEMLTNSVKELTS 378

Scoring table: BLOSUM62 Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

Database :   SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	AC	ID	Description
1	945	50.8	210	1	Q07337	OSC1_BORBU	borrelia bu
2	631	33.9	212	1	Q08137	OSC2_BORBU	borrelia bu
3	397.5	21.4	214	1	P32778	VM24_BORHE	borrelia he
4	397	21.4	215	1	P02448	VM03_BORHE	borrelia he
5	179	9.6	1509	1	P05559	MYSN_ACACA	acanthamoeb
6	173.5	9.3	1957	1	Q10411	YD86_SCHPO	schizosacch
7	168.5	9.1	483	1	P08009	M6_STRPY	streptococc
8	168.5	9.1	564	1	P19401	M12_STRPY	streptococc
9	167.5	9.0	998	1	Q9aix9	SCA4_RICAK	rickettsia
10	167.5	9.0	1251	1	Q0799	RBP2_PLAVB	plasmodium
11	166.5	8.9	1938	1	P24733	MYS_AEQIR	aequipecten
12	165.5	8.9	1790	1	P25386	USO1_YEAST	saccharomyc
13	164	8.8	2116	1	P08799	MYS2_DICDI	dictyosteli
14	163.5	8.8	3210	1	P49454	CENF_HUMAN	homo sapien
15	163	8.8	1966	1	P02566	MYSB_CAEEL	caenorhabdi
16	157.5	8.5	1786	1	P02469	LMB1_MOUSE	mus musculu
17	157	8.4	492	1	P02977	M5_STRPY	streptococc
18	155.5	8.4	1013	1	P0aj81	SCA4_RICRH	rickettsia
19	154	8.3	1427	1	P30622	REST_HUMAN	homo sapien
20	153.5	8.3	2869	1	Q00798	RBP1_PLAVB	plasmodium
21	153.5	8.3	1935	1	Q09339	MYSS_CYPCA	cyprinus ca
22	149.5	8.0	1011	1	Q9aj83	SCA4_RICAF	rickettsia
23	149	8.0	1189	1	Q90088	SCI1_CHICK	gallus gall
24	148.5	8.0	866	1	Q05870	MYSP_SCHJA	schistosoma
25	148	8.0	875	1	P31111	ZIP1_YEAST	saccharomyc
26	148	8.0	2017	1	Q99323	MYSN_DROME	drosophila
27	148	8.0	3110	1	P24063	LMA2_HUMAN	homo sapien
28	148	8.0	705	1	P34562	YNP9_CAEEL	caenorhabdi
29	147.5	7.9	1639	1	P15215	LMG1_DROME	drosophila
30	147	7.9	991	1	Q9ajd64	SCA4_RICAU	rickettsia
31	147	7.9	1500	1	P16952	SSP5_STRGN	streptococc
32	147	7.9	1961	1	Q62832	MYH9_RAT	rattus norv
33	146.5	7.9	967	1	P21613	KINH_LOLPE	loligo peal
34	146	7.9	679	1	P34237	YKR9_YEAST	saccharomyc
35	145.5	7.8	1786	1	P07942	LMB1_HUMAN	homo sapien
36	145	7.8	866	1	P06198	MYSP_SCHMA	schistosoma
37	145	7.8	1010	1	Q9aj75	SCA4_RICPA	rickettsia
38	144.5	7.8	1164	1	P27951	BAG_STRAG	streptococc
39	144.5	7.8	1433	1	042184	REST_CHICK	gallus gall
40	144.5	7.8	1453	1	Q9tu23	Y373_BOVIN	bos taurus
41	144.5	7.8	2349	1	P12270	TPR_HUMAN	homo sapien
42	144	7.7	1713	1	Q16787	LMA3_HUMAN	homo sapien
43	143.5	7.7	1935	1	P12883	MYH7_HUMAN	homo sapien
44	143.5	7.7	775	1	P33299	YIGE_BACSU	bacillus su
45	143	7.7	3672	1	Q21313	LML2_CAEEL	caenorhabdi

ALIGNMENTS

RESULT 1
OSC1_BORBU

ID   OSC1_BORBU   STANDARD;   PRT;   210 AA.
AC   Q07337;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   OUTER SURFACE PROTEIN C PRECURSOR (PC).
OS   Borrelia burgdorferi (Lyme disease spirochete).
OG   Plasmid lp54.
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 35210 / B31;
RX   MEDLINE=93289332; PubMed=8478108;
RA   Wilske B., Preac-Mursic V., Jauris S., Pradel I., Soutschek E.,
RA   Schwab E., Wanner G.;
RT   "Immunological and molecular polymorphisms of OspC, an immunodominant
RT   major outer surface protein of Borrelia burgdorferi.";
RL   Infect. Immun. 61:2182-2191(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 35210 / B31;
RX   MEDLINE=93268136; PubMed=8098841;
RA   Jauris-Heipke S., Fuchs R., Motz M., Preac-Mursic V., Schwab E.,
RA   Will G., Wilske B.;
RT   "Genetic heterogeneity of the genes coding for the outer surface
RT   protein C (OspC) and the flagellin of Borrelia burgdorferi.";
RL   Med. Microbiol. Immunol. 182:37-50(1993).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 35210 / B31;
RX   MEDLINE=94041630; PubMed=8225587;
RA   Padula S.J., Sampieri A., Dias F., Szczepanski A., Ryan R.W.;
RT   "Molecular characterization and expression of p23 (OspC) from a North
RT   American strain of Borrelia burgdorferi.";
RL   Infect. Immun. 61:5097-5105(1993).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 35210 / B31;
RX   MEDLINE=96025162; PubMed=7494039;
RA   Fukunaga M., Hamase A.;
RT   "Outer surface protein C gene sequence analysis of Borrelia
RT   burgdorferi sensu lato isolates from Japan.";
RL   J. Clin. Microbiol. 33:2415-2420(1995).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 35210 / B31;
RX   MEDLINE=98065943; PubMed=9403685;
RA   Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA   Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
RA   Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
RA   Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,
RA   van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,

RESULT 2

```
ID   OSC2_BORBU     STANDARD;      PRT;   212 AA.
AC   Q08137;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   OUTER SURFACE PROTEIN C PRECURSOR (PC).
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX   MEDLINE=92219995; PubMed=1560779;
RA   Fuchs R., Jauris S., Lottspeich F., Preac-Mursic V., Wilske B.,
```

```
RA   Utterback T., Watthey L., McDonald L., Artiach P., Bowman C.,
RA   Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
RA   Smith H.O., Venter J.C.;
RT   "Genomic sequence of a Lyme disease spirochaete, Borrelia
RT   burgdorferi.";
RL   Nature 390:580-586(1997).
CC   -!- FUNCTION: NOT KNOWN; MAJOR IMMUNODOMINANT PROTEIN.
CC   -!- SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID
CC       ANCHOR.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X69596; CAA49306.1; -.
DR   EMBL; U01894; AAA16058.1; -.
DR   EMBL; D64497; BAA08457.1; -.
DR   EMBL; AE000792; AAC66329.1; -.
DR   TIGR; BBB19; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
DR   PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KW   Outer membrane; Lipoprotein; Signal; Plasmid; Antigen;
KW   Complete proteome.
FT   SIGNAL        1     18       BY SIMILARITY.
FT   CHAIN        19    210       OUTER SURFACE PROTEIN C.
FT   LIPID        19     19       N-ACYL DIGLYCERIDE (BY SIMILARITY).
SQ   SEQUENCE   210 AA;  22340 MW;  7A4FC978F91777BF CRC64;

Query Match             50.8%;
Best Local Similarity   97.9%;
Matches  190;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEYEALLSSIDEIAAKA 60
              :|||||||||||||||||||||||||||||: ||||||||||||||:||||||||||||
Db         17 ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA 76

Qy         61 IGKKIHQNNGLDTENNNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 IGKKIHQNNGLDTENNNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN 136

Qy        121 KLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        137 KLKEKHTDLGKEGVTDADAKEAILKNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 196

Qy        181 ELTSPVVAESPKKP 194
              ||||||||||||||
Db        197 ELTSPVVAESPKKP 210
```

RESULT 3

```
ID   VM24_BORHE     STANDARD;      PRT;   214 AA.
AC   P32778;
DT   01-OCT-1993 (Rel. 27, Created)
DT   01-OCT-1993 (Rel. 27, Last sequence update)
DT   01-FEB-1994 (Rel. 28, Last annotation update)
DE   VARIABLE MAJOR OUTER MEMBRANE LIPOPROTEIN 24 PRECURSOR.
GN   VMP24.
OS   Borrelia hermsii.
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=140;
RN   [1]
RP   SEQUENCE FROM N.A.
```

```
RA   Soutschek E.;
RT   "Molecular analysis and expression of a Borrelia burgdorferi gene
RT   encoding a 22 kDa protein (pC) in Escherichia coli.";
RL   Mol. Microbiol. 6:503-509(1992).
RN   [2]
RP   SEQUENCE OF 1-205 FROM N.A.
RX   STRAIN=DK26;
RX   MEDLINE=94075528; PubMed=8253951;
RA   Theisen M., Frederiksen B., Lebech A.M., Vuust J., Hansen K.;
RT   "Polymorphism in OspC gene of Borrelia burgdorferi and
RT   immunoreactivity of OspC protein: implications for taxonomy and for
RT   use of OspC protein as a diagnostic antigen.";
RL   J. Clin. Microbiol. 31:2570-2576(1993).
CC   -!- FUNCTION: NOT KNOWN; MAJOR IMMUNODOMINANT PROTEIN.
CC   -!- SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID
CC       ANCHOR.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X62162; CAA44093.1; -.
DR   EMBL; X73624; CAA52003.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
KW   Outer membrane; Lipoprotein; Signal; Plasmid; Antigen.
FT   SIGNAL        1     18       BY SIMILARITY.
FT   CHAIN        19    212       OUTER SURFACE PROTEIN C.
FT   LIPID        19     19       N-ACYL DIGLYCERIDE (BY SIMILARITY).
SQ   SEQUENCE   212 AA;  22499 MW;  C206C231FBF2E7D4 CRC64;

Query Match             33.9%;
Best Local Similarity   69.9%;
Matches  137;  Conservative   15;  Mismatches   42;  Indels    2;  Gaps    2;

Qy          1 MACNNSGKDGNT-SANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAK 59
              :|||||| || : :|| |||| ||||||||| ||||||| ||||||||: ||||||: |
Db         17 ISCNNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETLVLSIDELAKK 76

Qy         60 AIGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKN-EGLKEKIDAAKKCSETF 118
              |||:||| :||| |  | :|||||||||||||| :||  ||| | |:|:|:||||||| |
Db         77 AIGQKIDNNNGLAALNNQNGSLLAGAYAISTLITEKLSKLKNLEELKTEIAKAKKCSEEF 136

Qy        119 TNKLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANS 178
              ||||| |  ||||: | |||:|||||| :  |  ||| | ||||||| ||||  ||  |
Db        137 TNKLKSGHADLGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGLLKAAQVALTNS 196

Qy        179 VKELTSPVVAESPKKP 194
              ||||||||||||||||
Db        197 VKELTSPVVAESPKKP 212
```

RESULT 4

ID   VM03_BORHE   STANDARD;   PRT;   215 AA.
AC   Q02448;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   01-FEB-1994 (Rel. 28, Last annotation update)
DE   VARIABLE MAJOR OUTER MEMBRANE LIPOPROTEIN 3 PRECURSOR.
GN   VMP3.
OS   Borrelia hermsii.
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=140;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SSP. HS1 SEROTYPE 3;
RX   MEDLINE=93133110; PubMed=1484486;
RA   Restrepo B.I., Kitten T., Carter C.J., Infante D., Barbour A.G.;
RT   "Subtelomeric expression regions of Borrelia hermsii linear plasmids
RT   are highly polymorphic.";
RL   Mol. Microbiol. 6:3299-3311(1992).
CC   -!- FUNCTION: SERVES TO AVOID THE HOST IMMUNE RESPONSE BY CHANGING
CC   FROM ONE SURFACE EXPOSED VMP TO ANOTHER.
CC   -!- SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID
CC   ANCHOR.
CC   -!- SIMILARITY: STRONG, TO VMP3.
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --
DR   EMBL; L04786; AAA22964.1; -.
DR   InterPro; IPR001800;.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   Prodom; PD001149; Lipoprotein_6;.
DR   PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KW   Outer membrane; Lipoprotein; Signal; Plasmid.
FT   SIGNAL   1   18
FT   CHAIN   19   214   VARIABLE MAJOR OUTER MEMBRANE
FT   LIPOPROTEIN 24.
FT   LIPID   19   19   N-ACYL DIGLYCERIDE (PROBABLE).
SQ   SEQUENCE   214 AA; 22541 MW; F1583F510246F7C7 CRC64;

Query Match   21.4%; Score 397.5; DB 1; Length 214;
Best Local Similarity   44.8%; Pred. No. 6.7e-12;
Matches   91; Conservative   37; Mismatches   58; Indels   17; Gaps   5;

Qy   1 MACNNSGKDGNTSANSADESVKGP----NLTEINKKITDSNAVLLAVKEVEALLSSIDEI 56
             ||:|||:|| :|| |:||| ::: :||:|||||:|||||:|||||:||:|:|:
Db   17 MSCNNGGPE----LKSDEVAKSDGTVLDLAKVSKKIKEASAFAASVKEVETLVKSVDEL 71

Qy   57 AAKAIGKKIHQNNGLDYENNHNGSLLAGAYAISTLIKQKL------DGLKNEGLKEKIDA 110
             |||||| ||| :| ||||| |||:|||:|:|| ::||| |||||| ||
Db   72 -AKAIGKKIKNDGGLDTEAGQNGSLIAGVHSVVSAVKIKVGALETTSGISNE-LKTKITE 129

Qy   111 AKKCSETFTNKLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKA 170
             |||| |:|||||||||| :||| |||| :| ||| |||| ||||:|||:|||||||||||
Db   130 VKSKAEAFLNKLKDGHTELGKKDASDDTKKAIKKDNSDKTKGASELEALNTAVDALLKA 189

Qy   171 AKKMLANSVKELTSPVVAESPKK 193
             ::||:|::||:||| |||:| :
Db   190 AEGEVEAAIKELTAPVKAEKPSQ 212

RESULT 5

ID   MYSN_ACACA   STANDARD;   PRT;   1509 AA.
AC   P05659;
DT   01-NOV-1988 (Rel. 09, Created)
DT   01-NOV-1988 (Rel. 09, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   MYOSIN II HEAVY CHAIN, NON MUSCLE.
OS   Acanthamoeba castellanii (Amoeba).
OC   Eukaryota; Acanthamoebidae; Acanthamoeba.
OX   NCBI_TaxID=5755;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87308395; PubMed=3040773;
RA   Hammer J.A. III, Bowers B., Paterson B.M., Korn E.D.;
RT   "Complete nucleotide sequence and deduced polypeptide sequence of a
RT   nonmuscle myosin heavy chain gene from Acanthamoeba: evidence of a
RT   hinge in the rodlike tail.";
RL   J. Cell Biol. 105:913-925(1987).
CC   -!- FUNCTION: MYOSIN IS A PROTEIN THAT BINDS TO F-ACTIN & HAS ATPASE
CC   ACTIVITY THAT IS ACTIVATED BY F-ACTIN.
CC   -!- SUBUNIT: MYOSIN II HEAVY CHAIN IS TWO-HEADED. IT SELF-ASSEMBLES
CC   INTO FILAMENTS. HEXAMER OF 2 HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI

Query Match   21.4%; Score 397; DB 1; Length 215;
Best Local Similarity   45.3%; Pred. No. 7.1e-12;
Matches   .92; Conservative   35; Mismatches   60; Indels   16; Gaps   5;

Qy   1 MACNNSGKDGNTSANSADESVKGP----NLTEINKKITDSNAVLLAVKEVEALLSSIDEI 56
             ||:|||:|| :|| |:|||::: :||:|||||:|||||:|||||:||:|:|:
Db   17 MSCNNGGPE----LKSDEVAKSDGTVLDLAKISKKIKDASDFAASVKEVHTLVKSIDEL 71

Qy   57 AAKAIGKKI-HQNNGLDYENNHNGSLLAGAYAISTLIKQKLDGLK------NEGLKEKIDA 110
             |||||||| :| ||| || ||||:||||:|:|| ::||| |||| |||||| ||
Db   72 -AKAIGKKIKNDNSNFEDENDHNGSLIAGVFQVILTVKAKLTSLEQIIGISDELKTEVGM 130

Qy   111 AKKCSETFTNKLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKA 170
             |||| |:||||||| :||| |||| :| ||| |||| ||||:|||:|||||||||||
Db   131 VKKESEAFVTQVKSKHTDLAKEGVTDAHAKSAILVTDGTKDKGAAELIKLNTAIDELLKA 190

Qy   171 AKEMLANSVKELTSPVVAESPKK 193
             :||:|:|::||:||| |||:|
Db   191 ANDAVETVIKELTASVKAEKPSQ 213

STRAIN=SSP. HS1 SEROTYPE 24;
RX   MEDLINE=93133110; PubMed=1484486;
RA   Restrepo B.I., Kitten T., Carter C.J., Infante D., Barbour A.G.;
RT   "Subtelomeric expression regions of Borrelia hermsii linear plasmids
RT   are highly polymorphic.";
RL   Mol. Microbiol. 6:3299-3311(1992).
CC   -!- FUNCTION: SERVES TO AVOID THE HOST IMMUNE RESPONSE BY CHANGING
CC   FROM ONE SURFACE EXPOSED VMP TO ANOTHER.
CC   -!- SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID
CC   ANCHOR.
CC   -!- SIMILARITY: STRONG, TO VMP24.
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --
DR   EMBL; L04789; AAA22967.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   Prodom; PD001149; Lipoprotein_6; 1.
DR   PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KW   Outer membrane; Lipoprotein; Signal; Plasmid.
FT   SIGNAL   1   18   PROBABLE.
FT   CHAIN   19   215   VARIABLE MAJOR OUTER MEMBRANE
FT   LIPOPROTEIN 3.
FT   LIPID   19   19   N-ACYL DIGLYCERIDE (PROBABLE).
SQ   SEQUENCE   215 AA; 23139 MW; 684C74D35F87C771 CRC64;

```
CC        LIGHT CHAIN SUBUNITS (MLC) AND 2 REGULATORY LIGHT CHAIN SUBUNITS
CC        (MLC-2).
-!-  DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CC        CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CC        CHARACTERISTIC OF ALPHA-HELICAL STRUCTURES. THIS REGION IS
CC        INTERRUPTED BY A HINGE AND JOINED BY A NONHELICAL TAILPIECE WHERE
CC        THE REGULATORY PHOSPHORYLATION SITES RESIDE.
-!-  MISCELLANEOUS: THE HINGE REGION MAY PLAY A KEY ROLE IN MEDIATING
CC        THE EFFECT OF HEAVY CHAIN PHOSPHORYLATION ON ENZYMATIC ACTIVITY.
-!-  SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
-!-  SIMILARITY: CONTAINS 2 IQ DOMAINS.
```

```
DR   EMBL; Y00624; CAA68663.1; -.
DR   PIR; A27224; A27224.
     HSSP; P08799; 1MND.
DR   InterPro; IPR000048; IQ.
DR   InterPro; IPR001609; myosin_head.
DR   Pfam; PF00612; IQ; 2.
DR   Pfam; PF00063; myosin_head; 1.
DR   PRINTS; PR00193; MYOSINHEAVY.
DR   ProDom; PD000355; myosin_head; 1.
DR   SMART; SM0001S; IQ; 1.
DR   SMART; SM00242; MYSc; 1.
     PROSITE; PS50096; IQ; 1.
KW   Myosin; Coiled coil; Actin-binding; ATP-binding; Calmodulin-binding;
KW   Methylation; Alkylation; Phosphorylation; Multigene family.
FT   DOMAIN      1    789    MYOSIN HEAD-LIKE.
FT   DOMAIN    790    819    IQ.
FT   DOMAIN    848   1509    COILED COIL (POTENTIAL).
FT   DOMAIN    848   1226    ALPHA-HELICAL TAILPIECE (S2).
FT   DOMAIN   1227   1252    HINGE.
FT   DOMAIN   1253   1509    LIGHT MEROMYOSIN (LMM).
FT   DOMAIN   1253   1482    ALPHA-HELICAL TAILPIECE (LMM).
FT   DOMAIN   1483   1509    NONHELICAL TAILPIECE.
FT   NP_BIND   182    189    ATP.
FT   DOMAIN    660    682    ACTIN-BINDING.
FT   DOMAIN    766    780    ACTIN-BINDING.
FT   DOMAIN    133    133    METHYLATION (TRI-) (POTENTIAL).
FT   MOD_RES   700    700    ALKYLATION (SH-1) (POTENTIAL).
FT   MOD_RES  1489   1489    PHOSPHORYLATION.
FT   MOD_RES  1494   1494    PHOSPHORYLATION.
FT   MOD_RES  1499   1499    PHOSPHORYLATION.
SQ   SEQUENCE  1509 AA;  171201 MW;  2CE49BE51173D17E CRC64;
```

```
Query Match             9.6%;
Best Local Similarity   24.9%;   Pred. No. 0.33;
Matches 102;  Conservative 69;  Mismatches 146;  Indels 92;  Gaps 20;
                        Score 179;  DB 1;  Length 1509;

Qy    28 EINKKITDSNAVLLAVKEVEA----LLSSIDEIAAKAIGKKIHQ--NNGLDTENNHNGSL 81
Db   934 ELQEEFSASNDILEQKRKLEAEKGELKASLEE--EERNRKALQEAKTVESERNELQDK 990

Qy    82 LAGAYAISTLIKQKLDGLKNEGLKEKIDA--AKKCSETFTNKLKEKHFDLGKEGVTDA- 137
Db   991 YEDEAAAHDSLKKKEEDLSRE-LRETKDALADAENISETFLRSKL--KNTERGADDVRNEL 1047

Qy   138 -DAKEAILKANGTKTKGAEELGKLFESVEVLISKAAKEMIANSVKELTSPV-------- 186
Db  1048 DDVTATKLQLEKTKKSLEEELAQTRAOLEE-EKSGKEAASSAKOLGQOLEDARSEVDSL 1106

Qy   187 ----VAESPRKPSMVNNSG----KDGNTSANSADESVKG--PNLTEISKKIT----ES 230
Db  1107 KSKLSAAEKSLKTAKDQNRDLDEQLEDERTVRANVDKQKKALEAKILFELEDQVTALDGQK 1166
```

RESULT   6
```
ID   YD86_SCHPO            STANDARD;      PRT;   1957 AA.
AC   Q10411;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   HYPOTHETICAL 222.8 KDA PROTEIN C1F3.06C IN CHROMOSOME I.
GN   SPAC1F3.06C.
OS   Schizosaccharomyces pombe (Fission yeast).
OC   Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC   Schizosaccharomycetales; Schizosaccharomycetaceae;
OC   Schizosaccharomyces.
OX   NCBI_TaxID=4896;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=972;
RA   Connor R., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;
RL   Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.
```

```
DR   EMBL; Z70690; CAA94624.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE  1957 AA;  222785 MW;  3F480CA061171D9DA CRC64;
```

```
Query Match             9.3%;
Best Local Similarity   22.8%;   Pred. No. 0.76;
Matches 108;  Conservative 73;  Mismatches 148;  Indels 145;  Gaps 18;
                        Score 173.5;  DB 1;  Length 1957;

Qy     8 KDGNTSANSADE--------SVKGPNLTEINKKITDSNAVLLAV-KEVEALLSSIDEI---- 56
Db   360 KDSRTSNSQLEEEMVELKESNRT-IHSQLTDAESKLSSFEQENKSLKGSIDEYQNNLSSK 418

Qy    57 -----AAKAIGKKIHONNGLDTEN----------------------NHNGSLL 82
Db   419 DKMVKQVSSQLEEARSSLAHAIGKLAEINSERDFQNKKIKDFEKIEODDLRACLNSSSNEL 478

Qy    83 AGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNKL--------KEKH---- 126
Db   479 KEK---SALIDKKDQELNN--LREQIKEQKVSESTQSSLQSLQRDILNEKKHHEVYESQ 533

Qy   127 ------TDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAA 171
Db   534 LNELKGELQTEISNSEHLSSQLSTLAAEKEAAVATN------NELSESKNSLQTICNAF 586

Qy   172 KEMIANSVKELTSPVVAESPKKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESN 231
Db   587 QEKLAKSVMQL----KENEQNFSSLDTSFKKLNESHOELEN-----NHQTITKQLKDTS 636

Qy   232 AVVLAVK-------------EVETLLTSIDELAKAIGKKITKNDVSLDNEADHNGS 273
Db   637 SKLQQLQLERANFEQKESTLSDENNDLRTKLKLEEESNKSLIKQEDVDSLE------- 688
```

Right-column alignment (Qy 231 region):
```
Qy   231 NAVVLAVKEVETLLITSIDEL----------AKAIGKKIKNDVSLDNEADHNGSLISGAYL 280
Db  1167 NA----AAAQAKTLKTQVDETKRRLEEAEASAARLEKERKN--ALDEVAQ------------ 1210

Qy   281 ISNLITKKISAIKDSG-------------ELKAEIEKAKK---CSEEFTAKLKGEHTDLGK 325
Db  1211 ----LTADLLDAERDSGAQORRKLNTRISELQSELENAPKTGGASSEEVKRLEGEELRLEE 1266

Qy   326 EGVTDDNAKKRAILKTNNDKTKGADELEKLFESVKNLSKAAKEMLTNSVK 374
Db  1267 ELLTAQEARAAAEK-NLDKAN--LELEELRQEADDAARDNDKLVKDNRK 1312
```

RESULT 7

```
Qy   274  LISGAYLISNLITKKISAIKDSGELKAEIEKAKKCSEEFTAKLKGEHTDLGKE-------G 327
               | |:| | | | | | | |:| | |:| |   |:|  |   |   | | |:|:||  |
Db   689  ------KNIQTLKEDLRKSEEALRESKLEAKNL-REVIDNLKGKHETLEAQRNDLHSS 739

Qy   328  VTDDNAKKAILKTNNDKTKGADELEKLPFESVKNL----SKAAKEMLTNSVKELTS 378
               ||:| ||  :|| | | |:||| |:::|| | | |   |  ||:||:||| | | |
Db   740  LSDAKNTNAIL--SSELTKSSEDVKRLTANVELTQDSKAMKQSFTSLVNSYQS 791
```

ID   M6_STRPY   STANDARD;   PRT;   483 AA.
AC   P0089;
DT   01-AUG-1988 (Rel. 08, Created)
DT   01-AUG-1988 (Rel. 08, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   M PROTEIN, SEROTYPE 6 PRECURSOR.
GN   EMM6.
OS   Streptococcus pyogenes.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC   Streptococcus.
OX   NCBI_TaxID=1314;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86111835; PubMed=3511046;
RA   Hollingshead S.K., Fischetti V.F., Scott J.R.;
RT   "Complete nucleotide sequence of type 6 M protein of the group A
RT   Streptococcus. Repetitive structure and membrane anchor.";
RL   J. Biol. Chem. 261:1677-1686(1986).
RN   [2]
RP   SEQUENCE OF 43-122 FROM N.A.
RX   MEDLINE=85166224; PubMed=3885219;
RA   Scott J.R., Pulliam W.M., Hollingshead S.K., Fischetti V.A.;
RT   "Relationship of M protein genes in group A streptococci.";
RL   Proc. Natl. Acad. Sci. U.S.A. 82:1822-1826(1985).
CC   -!- FUNCTION: THIS PROTEIN IS ONE OF THE DIFFERENT ANTIGENIC SEROTYPES
CC   OF PROTEIN M. PROTEIN M IS CLOSELY ASSOCIATED WITH VIRULENCE OF
CC   THE BACTERIUM AND CAN RENDER THE ORGANISM RESISTANT TO
CC   PHAGOCYTOSIS.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. CELL WALL.
CC   -!- SIMILARITY: TO OTHER M PROTEINS.
CC   -!- SIMILARITY: TO OTHER STREPTOCOCCAL AND STAPHYLOCOCCAL PROTEINS
CC   IN THE REGION OF THE MEMBRANE ANCHOR.
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --
DR   EMBL; M11338; AAA26920.1; -.
DR   PIR; A26297; A26297.
DR   InterPro; IPR001899; Gram_pos_anchor.
DR   InterPro; IPR003345; M_repeat.
DR   Pfam; PF00746; Gram_pos_anchor; 1.
DR   Pfam; PF02370; M; 9.
DR   PRINTS; PR00015; GPOSANCHOR.
DR   PROSITE; PS00343; GRAM_POS_ANCHORING; 1.
KW   Virulence; Phagocytosis; Cell wall; Duplication; Repeat; Antigen;
KW   Transmembrane; Coiled coil; Signal.
FT   SIGNAL   1   42
FT   CHAIN   43   483
FT   DOMAIN   43   457   M PROTEIN, SEROTYPE 6.
FT   DOMAIN   458   477   EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   478   483   MEMBRANE ANCHOR.
FT   DOMAIN   69   138   CYTOPLASMIC (POTENTIAL).
FT   DOMAIN   157   269   10 X 7 AA TANDEM REPEATS.
FT   DOMAIN   279   347   4.5 X 25 AA TANDEM REPEATS.
FT   DOMAIN   348   411   TWO DIRECTLY REPEATED 27 AMINO ACID
FT   BLOCKS SEPARATED BY 15 AMINO ACIDS.
FT   DOMAIN   412   448   GLY/PRO-RICH (CELL WALL-SPANNING).
FT   DOMAIN   449   454   CONSERVED IN GRAM-POSITIVE COCCI SURFACE
FT   PROTEINS.
SQ   SEQUENCE   483 AA;   53472 MW; 68F87F28DB53A448 CRC64;
```

```
Query Match            9.1%;   Score 168.5;  DB 1;  Length 483;
Best Local Similarity  22.9%;  Pred. No. 0.27;
Matches  91; Conservative 57; Mismatches 153;  Indels  97; Gaps 15;
```

```
Qy     5  NSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGKK 64
            |:|  |||: | |:| | |  | |:|  |:|| ||| |:|:|:|| |||:: |||: |
Db    73  NNDKLTTENNLITDQN---KNLTTENKNLTIDQNKNLTTENK----------------NL 112

Qy    65  IHQNGLLDGLKNEGLKEKIDAAKKCSEETFTNKLKE 124
            ||   | ||  |  |  :|:|   | :| |||
Db   214  TLKKTLDETVKDKIAKEQSKETI------GTLKKILDETVK----DKIARE-QKSKQD 261

Qy   113  TDQNKNLTTENK--ELKAEENRLTT----------ENKGLTKKLSEAE--EEAANKERE 157
            ||:| ||  | :|  | |:|  | |          |:|:| |:| :|  | |||||
Db   158  NKEAIGTLKKTLDETVKDKIAKEQESK----ETIGTLKKILDETVKDKIAKEQESKETIG 213

Qy   125  KHTDLGK--EGVTDADAKEAILKANGTKTKGAEELGKLFESVE--VLSKAAKEM------ 174
            |   | |  | |: | | |  |  |:| |      ||  |::|    |:||||:
Db   262  IGALKQELAKKDEGNNVSEASRGLRRDLDASREA------------------KRQVEK 302

Qy   175  -LANSVKELTSPVVAESPKKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAV 233
            :|||||: | | | | | | |:|: | | |: |:| | | | :|:  ||   |:| |||
Db   303  DLANLTAELDKVKEEKQISDASRQGLRRDLDASREAKKQVEKA-LEEANSKLAALEKLNK 361

Qy   234  VLAVKEVETLLTSIDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIK 293
            |||   | : | :|  |||  ||  :|   |:| | |: ||| ||:|||: ||| |||
Db   262  IGALKQELAKKDEGNNVSEASRGLRRDLDASREA------------------KRQVEK 302

Qy   294  DSGELKAEIEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEK 353
            |:| :|||| |:||:|| | ||  |  |:|:|   | | |  |:|:|||||||||:|| |
Db   303  DLANLTAELDKVKEEKQISDASRQGLRRDLDASREAKKQVEKA-LEEANSKLAALEKLNK 361

Qy   354  LFESVKNL------------------------SKAAKEMLTNSVKEL 376
            |||| |||                        |  ||||||||||||
Db   362  ELEESKKLTEKEKAELQAKLEAEAKALKEQIAEEL 399
```

RESULT 8

ID   M12_STRPY   STANDARD;   PRT;   564 AA.
AC   P19401;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   M PROTEIN, SEROTYPE 12 PRECURSOR (FRAGMENT).
GN   EMM12.
OS   Streptococcus pyogenes.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC   Streptococcus.
OX   NCBI_TaxID=1314;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CS24 / Serotype M12;
RX   MEDLINE=88058777; PubMed=2445730;
RA   Robbins J.C., Spanier J.G., Jones S.J., Simpson W.J., Cleary P.P.;
RT   "Streptococcus pyogenes type 12 M protein gene regulation by upstream
RT   sequences.";
RL   J. Bacteriol. 169:5633-5640(1987).
CC   -!- FUNCTION: THIS PROTEIN IS ONE OF THE DIFFERENT ANTIGENIC SEROTYPES
CC   OF PROTEIN M. PROTEIN M IS CLOSELY ASSOCIATED WITH VIRULENCE OF
CC   THE BACTERIUM AND CAN RENDER THE ORGANISM RESISTANT TO
CC   PHAGOCYTOSIS.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. CELL WALL.
CC   -!- SIMILARITY: TO OTHER M PROTEINS.
CC   -!- SIMILARITY: TO OTHER STREPTOCOCCAL AND STAPHYLOCOCCAL PROTEINS
CC   IN THE REGION OF THE MEMBRANE ANCHOR.
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way

RESULT 9

```
ID   SCA4_RICAK     STANDARD;      PRT;    998 AA.
AC   Q9AIX9;
DT   20-AUG-2001 (Rel. 40, Created)
DT   20-AUG-2001 (Rel. 40, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   ANTIGENIC HEAT-STABLE 120 KDA PROTEIN (PS120) (FRAGMENT)
DE   (PROTEIN PS 120) (120 KDA ANTIGEN)
GN   SCA4 OR D.
OS   Rickettsia akari.
```

```
CC   modified and this statement is not removed.  Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
DR   EMBL; M18269; AAA88573.1; -.
DR   PIR; A60115; A60115.
DR   InterPro; IPR001899; Gram_pos_anchor.
DR   InterPro; IPR003345; M_repeat.
DR   Pfam; PF00746; Gram_pos_anchor.
DR   Pfam; PF00746; Gram_pos_anchor; 1.
DR   Pfam; PF03370; M; 9
DR   PROSITE; PS00343; GRAM_POS_ANCHORING; 1.
KW   Virulence; Phagocytosis; Cell wall; Signal.
KW   Transmembrane; Coiled coil; Signal.
FT   SIGNAL        1     41
FT   CHAIN        42   >564
FT   DOMAIN       42    550       M PROTEIN, SEROTYPE 12.
FT   TRANSMEM    551   >564       MEMBRANE ANCHOR (POTENTIAL).
FT   DOMAIN       44    505       EXTRACELLULAR (POTENTIAL).
FT   DOMAIN      505    541       COILED COIL (POTENTIAL).
FT   DOMAIN      542    547       GLY/PRO-RICH (CELL WALL-SPANNING).
FT                               CONSERVED IN GRAM-POSITIVE COCCI SURFACE
FT                               PROTEINS.
SQ   SEQUENCE   564 AA;  62904 MW;  5F1549DACAA77B46 CRC64;
```

```
Query Match
Best Local Similarity   9.18%;
Conservative 22.9%;  Pred. No. 0.32;  DB 1;  Length 564;
Matches  97;  Conservative 59;  Mismatches 166;  Indels 101;  Gaps 17;
Score 168.5;

Qy   10  GNTSANSADESVKGPNLTEINKKIFDSNAVLLAV-KEVEALLSSIDEIAAKAIGKKIHON  68
             :||| |||::|||||||||||||||||||||||| |::||||::||||||||||::||
Db  115  GKLGIDNADLKAK----ITELEKSVEEKNDVLSQIKKELEEAEKDIQ------FGREVHAA 165

Qy   69  NGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNKLKEKHTD 128
Db  166  DLL------RIHKQEIAEKENVIS-----KLNG-ELQPLKQKVD------ETDRNLQQEKQKV 209

Qy  129  LGKE--GVTDADAKEAI-LKANGTKTKGAEELGKLFESVEVLSKAAK-EMLANSVKEL   182
Db  210  LSLEQQLAVTKENAKKDFELAALGHQLADKEYNAKIAELESKLADAAKDFELAALGHQHA 269

Qy  183  TSPVVAESPKKPSMVN----------NSGKDGNTSANSADESVKGPNLTEISK-----KI 227
Db  270  HNEYQAKLAEKDGQIKQLEEKQIQILDASRGTYARDLEAVRQAKKATEAELNNLKAELAKV 329

Qy  228  TESNAVVLAVK--------------EVETILTSIDE---LAKAIGKKIKNDVSLDNEA  268
Db  330  TEQKQILDASRKGTYARDLEAVRSKRQOOVEAALKQLEEQNKISEASRKGLRRDLDTSREA 389

Qy  269  DHNGSLISGAYLISNLITKKISAIKDSGELKAEIEKAKKCSEEFTAKLKGEHTDLGKEGV 328
Db  390  ------------KKQVEKDLANLTAELDKVKEEKQISDASRGLRRDLDASRE         430

Qy  329  TDDNAKKAILKTNNDKTKGADELEKLFESVKNL--------SKAAKEMLTNSV       373
Db  431  AKKQVEKA-LEEANSKLAALEKLNKDLEESKLTEKEAELQAKLEAEAKALKEOLAKOA   489

Qy  374  KEL 376
Db  490  EEL 492
```

RESULT 10

```
ID   RBP2_PLAVB     STANDARD;      PRT;   1251 AA.
AC   Q00799;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-APR-1993 (Rel. 25, Last sequence update)

OC   Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC   Rickettsiaceae; Rickettsieae; Rickettsia.
OX   NCBI_TaxID=786;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Sekeyova Z., Roux V., Raoult D.;
RT   "Phylogenetic analysis of Rickettsia spp. by comparing sequence of the
RT   'gene D' coding for an intracytoplasmic protein.";
RL   Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.
-!-  SUBCELLULAR LOCATION: CYTOPLASMIC (PROBABLE).
CC   -------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
DR   EMBL; AF213016; AAK30691.1; -.
KW   Antigen.
NON_TER       1       1
NON_TER       1       1
SQ   SEQUENCE   998 AA;  109328 MW;  FCEB43AC62DE5BD5 CRC64;
```

```
Query Match
Best Local Similarity   9.0%;
Conservative 20.7%;  Pred. No. 0.67;  DB 1;  Length 998;
Matches 102;  Conservative 69;  Mismatches 140;  Indels 181;  Gaps 22;
Score 167.5;

Qy    5  NSG-----KDGNTSANSADESV------KGPNLTEINKKIFDSNAVL-LAVKEVEAL     49
Db  424  NAGLTEKDGNTQIDLINEAATAILNNEKQANFITLIKNMVNNNALTPDTKVARVNAV    483

Qy   50  LSSIDEIAAKAIGKKIHQNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGL--KEK 107
Db  484  LETI----------KNNQDTPDIEKSKMLEATVAITI---------LNSENLTPKQK    520

Qy  108  IDAAKKCSETFTNKLKEKHTDLGKEGVTDADAKEAILKAN-GTKTKGAE--ELGKLFESV 164
Db  521  QQMLEKAVDVLSFKDDISRAVAIDGITG-----AVIKSNLSTKDGTMLIAVGDKVNAS    575

Qy  165  EVLSKAAKEMLANSV-----KELTSP--------VVAESPKKPSMVNNSGKDGNT      206
Db  576  E-LSNAEKQQLLGSVLKKGVVETFKILSPEQOQLMQQNLDKIIAEQTKNDNITEVQGTLANP 634

Qy  207  SANSADES-------------VKGPNLTEISKKITES---                     230
Db  635  AFNTIAKTPAAIOKVTTKVLDSPITAEIKGEFLLESTTKIVAESPLNVQDKTDIVKGMGEAI 694

Qy  231  ------NAVVLAVKEVET------LLTS------IDELAKAI                   254
Db  695  ASHRTMAPTKIKITAAIESVETGVAKSITDLEDKKLMTKGLVDGIYEDKANPEITSEMMKRAV 754

Qy  255  GKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAEIEKAKKCSEEFTA  314
Db  755  SKGVDN-----------STATPEDKQALKDAAS-EAALDRA---TQNFTE           789

Qy  315  KLKGEHTDLGKEGVTDDNAKKA------------------ILKTNNDKTKGADE--LEKLFESVK 359
Db  790  GLKGQNLDEPKP--RDDIYNKAQDIAYALKNVVTTVLDANPEKREVSEEEVMNKTSSILN  847

Qy  360  NLSKAAKEMLTN 371
Db  848  DISKIAIEKVNN 859
```

```
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   RETICULOCYTE BINDING PROTEIN 2 (FRAGMENT).
GN   RBP2.
OS   Plasmodium vivax (strain Belem).
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=31273;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92315338; PubMed=1617731;
RA   Galinski M.R., Medina C.C., Ingravallo P., Barnwell J.W.;
RT   "A reticulocyte-binding protein complex of Plasmodium vivax
RT   merozoites.";
RL   Cell 69:1213-1226(1992).
CC   -!- FUNCTION: INVOLVED IN RETICULOCYTE ADHESION. SPECIFICALLY BINDS TO
CC       HUMAN RETICULOCYTE CELLS.
CC   -!- SUBCELLULAR LOCATION: MEMBRANE-BOUND (PROBABLE).
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; M88098; AAA29744.1; -.
KW   Malaria; Receptor; Membrane.
NON_TER       1      1
NON_TER    1251   1251
SQ   SEQUENCE   1251 AA;  143741 MW;  54BA51C7404AC572 CRC64;

Query Match             9.0%;  Score 166.5;  DB 1;  Length 1251;
Best Local Similarity   22.68%;
Matches    91;  Conservative   55;  Pred. No. 0.95;  Mismatches 142;  Indels 115;  Gaps 17;

Qy     22 KGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGKKIHQNNGLGDTENNHNGSL 81
          |::|:|  |||  ||:|::|  |  :||  |   ::  | |
Db    774 KNTNELDVHKNIQDAVKVALEI------LAHSDEI-----------DTKQKDSSKL 812

Qy     82 LA--GAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNKL--KEKHTDLGKEGVT 135
Db    813 IEMGNQIYLKVLINQ------YKNKISSIKSKEEAVSVKIGNVSKKHSELSKITCS 863

Qy    136 DADAKEAI------LKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKELTSP 185
Db    864 DKSYDNIIALEKQTELQNLRNSFTQEKTNTNSDSKLEKI----KTDFESLKNALKTLEGE 919

Qy    186 VVAESPKKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLT 245
Db    920 VNA------LKASSDNHEHVQSKSEPV-NPALSEIEKEET--------DIDSLNT 959

Qy    246 SIDELAK----AIGKKIKNDVSLDNEADHNGSLISGAYL------AYLAYIKKNYEDT 280
Db    960 ALDELLKKGRTCEVSRYKLIKDTVTKEISDDTELINTIEKNVK-------- 1015

Qy    281 ISNLIT--KKISAIKDSGELKAEIEKAKKCSEEFTARLKGEHTDLGKEGVTDDNAKAIL 338
Db   1016 VQDVLTLNEHFNTKQVSNHEPTNFDKSNKSSEELTKAVTDSKTIISK------LKGVII 1068

Qy    339 KTN-----NDKTKGADELEKLFESVKNLSKAAKEM--LTNSVK 374
Db   1069 EVNENTEMNTIESSAKIEALYNELKNKKTSLNEIYQTSNEVK 1111
```

```
RESULT 11
ID   MYS_AEQIR     STANDARD;      PRT;  1938 AA.
AC   P24733;
DT   01-MAR-1992 (Rel. 21, Created)
DT   01-MAR-1992 (Rel. 21, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   MYOSIN HEAVY CHAIN, STRIATED MUSCLE.
OS   Aequipecten irradians (Bay scallop).
OC   Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorphia; Pectinoida;
OC   Pectinoidea; Pectinidae; Argopecten.
OX   NCBI_TaxID=31199;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Adductor muscle;
RX   MEDLINE=92011595; PubMed=1917970;
RA   Nyitray L., Goodwin E.B., Szent-Gyorgyi A.G.;
RT   "Complete primary structure of a scallop striated muscle myosin heavy
RT   chain. Sequence comparison with other heavy chains reveals regions
RT   that might be critical for regulation.";
RL   J. Biol. Chem. 266:18469-18476(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Adductor muscle;
RX   MEDLINE=91088319; PubMed=2263488;
RA   Nyitray L., Goodwin E.B., Szent-Gyorgyi A.G.;
RT   "Nucleotide sequence of full length cDNA for a scallop striated
RT   muscle myosin heavy chain.";
RL   Nucleic Acids Res. 18:7158-7158(1990).
RN   [3]
RP   X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 777-836.
RX   MEDLINE=94173332; PubMed=8127365;
RA   Xie X., Harrison D.H., Schlichting I., Sweet R.M., Kalabokis V.N.,
RA   Szent-Gyorgyi A.G., Cohen C.;
RT   "Structure of the regulatory domain of scallop myosin at 2.8-A
RT   resolution.";
RL   Nature 368:306-312(1994).
RN   [4]
RP   X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 777-836.
RX   MEDLINE=96419133; PubMed=8805510;
RA   Houdusse A., Cohen C.;
RT   "Structure of the regulatory domain of scallop myosin at 2-A
RT   resolution: implications for regulation.";
RL   Structure 4:21-32(1996).
CC   -!- FUNCTION: MUSCLE CONTRACTION.
CC   -!- FUNCTION: MYOSIN IS A PROTEIN THAT BINDS TO F-ACTIN & HAS ATPASE
CC       ACTIVITY THAT IS ACTIVATED BY F-ACTIN.
CC   -!- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
CC       HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
CC       AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
CC   -!- SUBCELLULAR LOCATION: THICK FILAMENTS OF THE MYOFIBRILS.
CC   -!- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 IQ DOMAIN.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X55714; CAA39247.1; -.
DR   PIR; S13557; S13557.
DR   PIR; A40997; A40997.
DR   PDB; 1SCM; 30-APR-94.
DR   PDB; 1WDC; 11-JUL-96.
DR   InterPro; IPR000048; IQ.
DR   InterPro; IPR002928; Myosin_tail.
DR   InterPro; IPR002017; Spectrin.
DR   InterPro; IPR001609; Myosin_head.
DR   Pfam; PF00612; IQ; 1.
DR   Pfam; PF00063; myosin_head; 1.
DR   Pfam; PF01576; Myosin_tail; 1.
DR   PRINTS; PR00193; MYOSINHEAVY.
DR   ProDom; PD000355; myosin_head; 1.
DR   SMART; SM00015; IQ; 1.
DR   SMART; SM00242; MYc; 1.
DR   PROSITE; PS50096; IQ; 1.
KW   Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
KW   ATP-binding; Alkylation; Calmodulin-binding; 3D-structure.
```

FT   DOMAIN   1   777   MYOSIN HEAD-LIKE.
FT   DOMAIN   778   805   IQ.
FT   DOMAIN   836   1938   RODLIKE TAIL (S2 AND LMM DOMAINS).
FT   DOMAIN   836   1938   COILED COIL (POTENTIAL).
FT   NP_BIND   176   183   ATP (BY SIMILARITY).
FT   MOD_RES   693   693   ALKYLATION (SH-1) (BY SIMILARITY).
FT   MOD_RES   703   703   ALKYLATION (SH-2) (BY SIMILARITY).
FT   HELIX   778   821
FT   HELIX   822   823
FT   TURN   822   823
FT   HELIX   825   833
FT   TURN   834   835
FT   SEQUENCE   1938 AA; 222821 MW; A5CCE4127D1A4896 CRC64;

Query Match   8.9%; Score 165.5; DB 1; Length 1938;
Best Local Similarity 22.3%; Pred. No. 1.7;
Matches   88; Conservative 59; Mismatches 167; Indels 81; Gaps 12;

```
Qy     18 DESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGKKIHQNNGLDT----- 73
             ||| |: : :|: :| |||  :|||: |: |: : | :: | ::|: :|
Db    925 DEEDAAADLEGIKKKMEADNA-------NLKKDIGDLENTLQKAEQDKAHKDNQISTLQGE 978

Qy     74 ---ENNHNGSLLAGAYAISTLIKOKLDGLKNEGLKEKIDAAKKCSEFTFNKLKEKHTDLG 130
            :  :|| | | | |: :| | :| ::|||||||||||| : ::|||
Db    979 ISQODEHIGKLNKEKKALEEANKKTSDSLQAE--EDKCNLHLNKLKAKLEQALDELEDNLE 1036

Qy    131 KEGVTDADAKEALLKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKELTSPVVAES 190
            :| :| |||  |:  : |::
Db   1037 REKKVRGDVEKA------KRKVEQDLKSTQENVEDLERVKRELEEN----------VRRK 1080

Qy    191 PKKPSMVNNSGKDGNTSANSADESVK--GPNLTEISKKI-TESNAVVLAVKEVETLLTSI 247
            :|: : : :| ::| :
Db   1081 EAEISSLNSKLEDEQNLVSQLQRIKELQARIEELEEELEAERNARAKVEKQRAELNREL 1140

Qy    248 DELAKAIGK--------------KIKNDV---SLDNEADHNGSLISGAYLISNL 284
            :|| | : || || |||
Db   1141 EELGERLDEAGGATSAQIELNKKREAELLKIRRDLEEASLQHEAQ---------ISAL 1189

Qy    285 ITKKISAIKDSGELKAEIEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDK 344
            |: |:|| :| : :| || |||
Db   1190 RKKHQDAANEMADQVDQLQKVKSKLEKDKIDLKREMDDLESQ----------MTHNMK 1237

Qy    345 TKGADE-LEKLFESVKNLSKAAKEMLTNSVKELTS 378
            :| | : :| | |: :|
Db   1238 NKGCSEKVMKQFESQMSDLNARLEDSQRSINELQS 1272
```

SLT 12

ID   USO1_YEAST   STANDARD;   PRT;   1790 AA.
AC   P25386;
DT   01-MAY-1992 (Rel. 22, Created)
DT   01-MAY-1992 (Rel. 22, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   INTRACELLULAR PROTEIN TRANSPORT PROTEIN USO1.
GN   USO1 OR INT1 OR YDL058W.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=X2180-1A;
RX   MEDLINE=91185402; PubMed=2010462;
RA   Nakajima H., Hirata A., Ogawa Y., Yonehara T., Yoda K.,
RA   Yamasaki M.;
RT   "A cytoskeleton-related gene, uso1, is required for intracellular
RT   protein transport in Saccharomyces cerevisiae.";
RL   J. Cell Biol. 113:245-260(1991).
RN   [2]
RP   SEQUENCE OF 782-1790 FROM N.A.
RA   Hostetter M.K., Herman D.J., Bendel C.M., McClellan M., Tao N.,
RA   Kendrick K.E.;

RL   Submitted (FEB-1993) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 1-8 FROM N.A.
RA   Bai Y., Symington L.S.;
RL   Submitted (MAY-1996) to the EMBL/GenBank/DDBJ databases.
-!- FUNCTION: REQUIRED FOR PROTEIN TRANSPORT FROM THE ER TO THE GOLGI
    COMPLEX.
-!- SUBCELLULAR LOCATION: CYTOPLASMIC. ASSOCIATED WITH INTRACELLULAR
    MEMBRANES. PROBABLY PRESENT ON VESICLES OPERATIONAL BETWEEN THE
    ER AND THE GOLGI COMPLEX.
-!- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, COMPOSED
    OF AN HEPTAPEPTIDE REPEAT PATTERN CHARACTERISTIC OF ALPHA-HELICAL
    COILED COILS. MAY FORM FILAMENTOUS STRUCTURES IN THE CELL.
-!- SIMILARITY: BELONGS TO THE VDP/USO1/YBL047C FAMILY.

DR   EMBL; X54378; CAA38253.1; -.
DR   EMBL; L03188; AAB00143.1; -.
DR   EMBL; U53668; AAB66659.1; -.
DR   PIR; A38455; A38455.
DR   HSSP; P80220; 1DIP.
DR   SGD; S0002216; USO1.
DR   InterPro; IPR002017; Spectrin.
KW   Transport; Protein transport; Golgi stack; Cytoskeleton; Coiled coil.
FT   DOMAIN   1   724   GLOBULAR HEAD.
FT   DOMAIN   725   1790   COILED COIL (POTENTIAL).
FT   DOMAIN   465   487   CHARGED (HYPER-HYDROPHILIC).
FT   DOMAIN   991   1172   DISPENSABLE FOR THE PROTEIN FUNCTION.
FT   DOMAIN   1172   1786   ASP/GLU-RICH (ACIDIC).
FT   CONFLICT   847   847   G -> E (IN REF. 2).
FT   CONFLICT   924   924   G -> K (IN REF. 2).
FT   CONFLICT   1253   1253   V -> I (IN REF. 2).
FT   CONFLICT   1319   1319   V -> V (IN REF. 2).
FT   CONFLICT   1461   1461   N -> S (IN REF. 2).
FT   CONFLICT   1581   1581   G -> S (IN REF. 2).
FT   CONFLICT   1600   1600   G -> S (IN REF. 2).
FT   CONFLICT   1661   1661   R -> S (IN REF. 2).
FT   CONFLICT   1772   1772   D -> DEEDDEE (IN REF. 2).
FT   SEQUENCE   1790 AA; 206424 MW; 6CE2B2169FD4818 CRC64;

Query Match   8.8%; Score 164; DB 1; Length 1790;
Best Local Similarity 21.6%; Pred. No. 1.8;
Matches 101; Conservative 75; Mismatches 186; Indels 106; Gaps 17;

```
Qy      3 CNNSGKDG-----------NTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSS 52
           ||| |:
Db    929 CNNLSKEKEHISKELVEYKSRFQSHDNLV--AKLTEKLKSLANNYKDMQA--ENESLIKA 984

Qy     53 IDEIAAKAIGKKIHQNNGLD--TENNHNGSLLAGAY--------AISTL-----IKQK 95
           :|: :| | | :: |
Db    985 VEESKNESSIQLSNLQNKIDSMSQEKENFQIERGSIEKNIEQLIKKITSDLEQTKEEIISK 1044

Qy     96 LDGIKNEG---------LKEKIDAAKKCSETFTNKLKE--------KHTDLGKEGVTD 136
           | |:|: ||||:|:|: :|: :|: | | ||
Db   1045 SDSSKDYESQISLLKEKLETATTANDENVNKISELTKTREELEAELAAYKNLKNELETK 1104

Qy    137 ADAKEAILK---------ANGTKTK------GAEELGKLFESVEVLSKAA 171
           | | | :|||||||||
Db   1105 LETSEKALKEYKENEEHLKEEKIQLEKEATETKQOLNSLRANLESLEKEHEDLAAQLKKY 1164

Qy    172 KEMLANSVKELTSPV--VAESPKKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITE 229
           :|: :|::| |: :| : :| :::
Db   1165 EEQIANKERQINEEISQLNDEITSTQOENESIKKKNDELEGEVKAMKSTSEEQSNLKKSE 1224

Qy    230 SNAVVLAVKEV----ETLLTSIDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLI 285
```

```
Db  1225 IDALNLQIKELKKKNETNEASLLESIKSVESETVKIRELQDECNFKEKEVSE------L 1277
          ::|:|| :||:| :|| :|||
Qy   286 TKKISAIKDSG----ELKAEIEKAKCSEEFTAKLK---GEHIDLGKEGVTDDNAKKAIL 338
          : | || :|| |: | |:| :| :|
Db  1278 EDKLKASEDKNSKYLELQKESEKIKELDAKTTELKIQLEKIINLSKAKEKSESELSRLK 1337
          | :|| |||   |:| :|||
Qy   339 KTNNDKTKGADE-LEKL---------FESVKNLSKAAKEMLTNSVKE 375
          | :| |||    |:|
Db  1338 KTSSEERKNAEEQLELKIKNEIQIKNQAFEKRKLLNEGSSTITQEYSE 1385
```

```
RESULT 13
ID   MYS2_DICDI     STANDARD;      PRT;   2116 AA.
AC   P08799;
DT   01-NOV-1988 (Rel. 09, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   MYOSIN II HEAVY CHAIN, NON MUSCLE.
GN   MHCA.
OS   Dictyostelium discoideum (Slime mold).
OC   Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX   NCBI_TaxID=44689;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87092266; PubMed=3540939;
RA   Warrick H.M., de Lozanne A., Leinwand L.A., Spudich J.A.;
RT   "Conserved protein domains in a myosin heavy chain gene from
RT   Dictyostelium discoideum.";
RL   Proc. Natl. Acad. Sci. U.S.A. 83:9433-9437(1986).
RN   [2]
RP   PHOSPHORYLATION SITES, AND MUTAGENESIS.
RC   STRAIN=AX2;
RX   MEDLINE=88122226; PubMed=2828113;
RA   Wagle G., Noegel A., Scheel J., Gerisch G.;
RT   "Phosphorylation of threonine residues on cloned fragments of the
RT   Dictyostelium myosin heavy chain.";
RL   FEBS Lett. 227:71-75(1988).
RN   [3]
RX   MEDLINE=90353583; PubMed=2387408;
RA   Lueck-Vielmeter D., Schleicher M., Grabatin B., Wippler J.,
RA   Gerisch G.;
RT   "Replacement of threonine residues by serine and alanine in a
RT   phosphorylatable heavy chain fragment of Dictyostelium myosin II.";
RL   FEBS Lett. 269:239-243(1990).
RN   [4]
RP   X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 1-762.
RX   MEDLINE=95345066; PubMed=7619795;
RA   Fisher A.J., Smith C.A., Thoden J.B., Smith R., Sutoh K., Holden H.M.,
RA   Rayment I.;
RT   "X-ray structures of the myosin motor domain of Dictyostelium
RT   discoideum complexed with MgADP.BeFx and MgADP.AlF4-.";
RL   Biochemistry 34:8960-8972(1995).
RN   [5]
RP   X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 1-762.
RX   MEDLINE=95345067; PubMed=7619796;
RA   Smith C.A., Rayment I.;
RT   "X-ray structures of the magnesium(II)-pyrophosphate complex of the
RT   truncated head of Dictyostelium discoideum myosin to 2.7-A
RT   resolution.";
RL   Biochemistry 34:8973-8981(1995).
RN   [6]
RP   X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 1-762.
RX   MEDLINE=96206189; PubMed=8611530;
RA   Smith C.A., Rayment I.;
RT   "X-ray structure of the magnesium(II).ADP.vanadate complex of the
RT   Dictyostelium discoideum myosin motor domain to 1.9-A resolution.";
RL   Biochemistry 35:5404-5417(1996).
RN   [7]
RX   MEDLINE=97452580; PubMed=9305951;
```

```
RA   Gulick A.M., Bauer C.B., Thoden J.B., Rayment I.;
RT   "X-ray structures of the MgADP, MgATPgammaS, and MgAMPPNP complexes
RT   of the Dictyostelium discoideum myosin motor domain.";
RL   Biochemistry 36:11619-11628(1997).
RN   [8]
RP   X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 1-762.
RX   MEDLINE=98070605; PubMed=9405148;
RA   Bauer C.B., Kuhlman P.A., Bagshaw C.R., Rayment I.;
RT   "X-ray crystal structure and solution fluorescence characterization
RT   of Mg.2'(3')-O-(N-methylanthraniloyl) nucleotides bound to the
RT   Dictyostelium discoideum myosin motor domain.";
RL   J. Mol. Biol. 274:394-407(1997).
CC   -!- FUNCTION: MYOSIN IS A PROTEIN THAT BINDS TO ACTIN & HAS ATPASE
CC       ACTIVITY THAT IS ACTIVATED BY ACTIN.
CC   -!- SUBUNIT: MYOSIN II HEAVY CHAIN IS TWO-HEADED. IT SELF-ASSEMBLES
CC       INTO FILAMENTS. HEXAMER OF 2 HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI
CC       LIGHT CHAIN SUBUNITS (MLC) AND 2 REGULATORY LIGHT CHAIN SUBUNITS
CC       (MLC-2).
CC   -!- SUBCELLULAR LOCATION: HIGHEST CONCENTRATION IN THE POSTERIOR CELL
CC       CORTEX.
CC   -!- DOMAIN: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
CC       MEROMYOSIN (LMM) AND 1 HEAVY MEROMYOSIN (HMM). IT CAN BE FURTHER
CC       SPLIT INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
CC       SUBFRAGMENT (S2).
CC   -!- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CC       CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CC       CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC   -!- PTM: PHOSPHORYLATION INHIBITS THICK FILAMENT FORMATION AND REDUCES
CC       THE ACTIN-ACTIVATED ATPASE ACTIVITY.
CC   -!- MISCELLANEOUS: DICTYOSTELIUM MYOSIN II HAS NO K(2)EDTA ATPASE
CC       ACTIVITY, PERHAPS CORRELATED WITH THE ABSENCE OF A CYS AT THE SH-1
CC       POSITION (688).
CC   -!- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 IQ DOMAIN.
```

```
DR   EMBL; M14628; AAA33227.1; -.
DR   PIR; A26655; A26655.
DR   PIR; S00250; S00250.
DR   PDB; 1MMA; 03-DEC-97.
DR   PDB; 1MMD; 17-AUG-96.
DR   PDB; 1MMG; 03-DEC-97.
DR   PDB; 1MMN; 03-DEC-97.
DR   PDB; 1MND; 17-AUG-96.
DR   PDB; 1MNE; 17-AUG-96.
DR   PDB; 1VOM; 23-DEC-96.
DR   PDB; 1LVK; 28-JAN-98.
DR   DictyDb; DD01008; mhcA.
DR   InterPro; IPR000048; IQ.
DR   InterPro; IPR001609; myosin_head.
DR   Pfam; PF00612; IQ; 1.
DR   Pfam; PF00063; myosin_head; 1.
DR   PRINTS; PR00193; MYOSINHEAVY.
DR   ProDom; PD000355; myosin_head; 1.
DR   SMART; SM00015; IQ; 1.
DR   SMART; SM00242; MYSc; 1.
DR   PROSITE; PS50096; IQ; 1.
KW   Myosin; Coiled coil; Actin-binding; ATP-binding; 3d-structure;
KW   Calmodulin-binding; Methylation; Alkylation; Phosphorylation.
FT   DOMAIN        1    761       MYOSIN HEAD-LIKE.
FT   DOMAIN      762    791       IQ.
FT   DOMAIN      817   2116       COILED COIL (POTENTIAL).
FT   NP_BIND     179    186       ATP.
FT   DOMAIN      638    660       ACTIN-BINDING.
FT   DOMAIN      738    752       ACTIN-BINDING.
FT   MOD_RES     130    130       METHYLATION (DI-) (POTENTIAL).
```

```
FT   MOD_RES       678    678    ALKYLATION (SH-1).
FT   MOD_RES      1823   1823    PHOSPHORYLATION (BY MHCK).
FT   MOD_RES      1833   1833    PHOSPHORYLATION (BY MHCK).
FT   MOD_RES      2029   2029    PHOSPHORYLATION (BY MHCK).
SQ   SEQUENCE   2116 AA;  243871 MW;  2FC3770B1EE56A1 CRC64;
```

Query Match   8.8%; Score 163.5; DB 1; Length 2116;
Best Local Similarity 21.0%; Pred. No. 2.3;
Matches   98; Conservative   73; Mismatches 152; Indels 143; Gaps 18;

```
Qy     9  DGNTSANSADESVKG------PNLIELNKKITDSNAVLLAVKEVEALLSSIDEIAKAIG   62
             ||:|        ||     |:|: |||::   |:|: :: |:   |  : |::|: | 
Db   964  DGQSDTISRLEKIKDELQKEVEELIESFSEESKDKGVL--EKTRVRLQSELDDLTVR---  1018

Qy    63  KKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSEFETNKL  122
             |:|              |:  | ||: |||:|  |: |: |:::  :     |:  |:
Db  1019  LDSETKDKSELLRQKKKKLEEELKQVQEALAAE------TAAKLAQEAANKKL        1064

Qy   123  KEKHTDLGKEGVTDADAKEAILKANGT----------KTKGA----               154
             |:| |: |: | : : | | : ||              ||
Db  1065  QGEYTELNEKPNSEVTARSNVEKSKKTLESQLVAVNNELEDEKKNRDALEKKKALDAML  1124

Qy   155  EELGKLFES--------VEVLSKAAKEMLANSVKELTSPVV--------             187
             ||: |:        :| :|:| |:::: | :::||
Db  1125  EEMKDQLESTGGEKKSIYDLKVKQESDMEALRNQISELQSTIAKLEKIKSTLEGEVARLQ  1184

Qy   188  ------AESPKKPSMVNNSGKDGNTSANSADESVKGPNLTEISKRITESNAVVL        235
               :|::| | : :  :|:|| :||::  | |:| :  ::  ::::  :  | 
Db  1185  GELEAEQLAKSNVEKQKKKVELDLEDK----SAQLAEETAAKQALDKLKKKLEQE---   1235

Qy   236  AVKEVETLLTSIDELAKAIGKKIK----------NDVSLDNEADHNGSLISGAVLISN    283
              ::|::  ::|:|  |:| :        ||       || :| | |:| 
Db  1236  -LSEVQT-----QLSEANKNVNSDSTNKHLETSPNNLKLELEAEQKAK-------QA     1280

Qy   284  LITKKI---SAIKDSSGELKAEIEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNA-KKAILK  339
             |::|:   | |: :| ::: |::: :: :  |: |: ||  |: |: | :::| :||| 
Db  1281  LEKKRLGLESELKHVNEQLEEEKKQKESNEKRKVDLEKEVSEL-KQIEEVASKKAVTE     1339

Qy   340  TNNDKTKGADELEKLF-----------ESVKNLS--KAAKEMLTNSVKE             375
              :| |:| |:| ||:|           ::|:|||  |||||:| ||||:
Db  1340  AKNKKESELDEIKRQYADVVSSRDKSVEQLKTLQAKNEELRNTAEE                1385
```

RESULT 14
CENF_HUMAN   STANDARD;   PRT;   3210 AA.

```
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   CENP-F KINETOCHORE PROTEIN (CENTROMERE PROTEIN F) (MITOSIN) (AH
DE   ANTIGEN).
GN   CENPF.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Breast carcinoma;
RX   MEDLINE=95348175; PubMed=7542657;
RA   Liao H., Winkfein R.J., Mack G., Rattner J.B., Yen T.J.;
RT   "CENP-F is a protein of the nuclear matrix that assembles onto
RT   kinetochores at late G2 and is rapidly degraded after mitosis.";
RL   J. Cell Biol. 130:507-518(1995).
RN   [2]
RX   MEDLINE=9579848; PubMed=7651420;
RA   Zhu X., Mancini M.A., Chang K.-H., Liu C.-Y., Chen C.-F., Shan B.,
RA   Jones D., Yang-Feng T.L., Lee W.-H.;
RT   "Characterization of a novel 350-kilodalton nuclear phosphoprotein
RT   that is specifically involved in mitotic-phase nuclear progression.";
```

```
RL   Mol. Cell. Biol. 15:5017-5029(1995).
RN   [3]
RP   SEQUENCE OF 2194-3210 FROM N.A.
RX   MEDLINE=95356446; PubMed=7612011;
RA   Zhu X., Chang K.-H., He D., Mancini M.A., Brinkley W.R., Lee W.-H.;
RT   "The C terminus of mitosin is essential for its nuclear localization,
RT   centromere/kinetochore targeting, and dimerization.";
RL   J. Biol. Chem. 270:19545-19550(1995).
RN   [4]
RX   MEDLINE=95370296; PubMed=7642639;
RA   Li Q., Ke Y., Kapp J.A., Fertig N., Medsger T.A. Jr., Joshi H.C.;
RT   "A novel cell-cycle-dependent 350-kDa nuclear protein; C-terminal
RT   domain sufficient for nuclear localization.";
RL   Biochem. Biophys. Res. Commun. 212:220-228(1995).
RN   [5]
RP   CHARACTERIZATION.
RX   MEDLINE=98437347; PubMed=9763420;
RA   Chan G.K.T., Schaar B.T., Yen T.J.;
RT   "Characterization of the kinetochore binding domain of CENP-E reveals
RT   interactions with the kinetochore proteins CENP-F and hBUBR1.";
RL   J. Cell Biol. 143:49-63(1998).
CC   -!- FUNCTION: PROBABLY REQUIRED FOR KINETOCHORE FUNCTION, INVOLVED IN
CC       CHROMOSOME SEGREGATION DURING MITOSIS. INTERACTS WITH
CC       RETINOBLASTOMA PROTEIN (RB), CENP-E AND BUB1.
CC   -!- SUBUNIT: HOMO- OR HETERODIMER.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR MATRIX (BUT NOT IN THE NUCLEOLUS),
CC       REORGANIZATION TO THE KINETOCHORE/CENTROMERE (CORONAL SURFACE OF
CC       THE OUTER PLATE) AND THE SPINDLE DURING MITOSIS.
CC   -!- DEVELOPMENTAL STAGE: GRADUALLY ACCUMULATES DURING THE CELL CYCLE.
CC   -!- PTM: HYPERPHOSPHORYLATED DURING MITOSIS.
CC   -!- ------------------------------------------------------------------
CC       This SWISS-PROT entry is copyright. It is produced through a collaboration
CC       between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC       the European Bioinformatics Institute. There are no restrictions on its
CC       use by non-profit institutions as long as its content is in no way
CC       modified and this statement is not removed. Usage by and for commercial
CC       entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC       or send an email to license@isb-sib.ch).
CC   -!- ------------------------------------------------------------------
DR   EMBL; U19769; AAA82889.1; -.
DR   EMBL; U30872; AAA82935.1; -.
DR   EMBL; U25725; AAA86889.1; -.
DR   HSSP; P02649; 1LE4.
DR   MIM; 600236; -.
KW   Chromosomal protein; Nuclear protein; Centromere; Coiled coil;
KW   Mitosis; Phosphorylation; Antigen; Cell cycle; Repeat.
FT   DOMAIN         14     197    COILED COIL (POTENTIAL).
FT   DOMAIN        273     769    COILED COIL (POTENTIAL).
FT   DOMAIN        823    1328    COILED COIL (POTENTIAL).
FT   DOMAIN       1642    1746    COILED COIL (POTENTIAL).
FT   DOMAIN       1862    2987    COILED COIL (POTENTIAL).
FT   DOMAIN       2207    2568    COILED COIL (POTENTIAL).
FT   REPEAT       2207    2386    2 X 177 AA TANDEM REPEATS.
FT   REPEAT       2389    2568    2.
FT   SITE         3015    3032    NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT   CONFLICT       16      16    T -> A (IN REF. 2).
FT   CONFLICT      250     250    L -> V (IN REF. 2).
FT   CONFLICT      272     272    G -> D (IN REF. 2).
FT   CONFLICT      611     611    MISSING (IN REF. 2).
FT   CONFLICT     1494    1589    MISSING (IN REF. 2).
FT   CONFLICT     1611    1611    V -> A (IN REF. 2).
FT   CONFLICT     1811    1811    V -> L (IN REF. 2).
FT   CONFLICT     2242    2243    ER -> DG (IN REF. 3).
FT   CONFLICT     2335    2335    L -> Q (IN REF. 3).
FT   CONFLICT     2492    2492    D -> N (IN REF. 2).
FT   CONFLICT     2545    2561    ELNERVAALHNDQEACK -> SSMREWQPCIMTKKPVS
FT                               (IN REF. 3).
SQ   SEQUENCE   3210 AA;  367589 MW;  11D8332496064334 CRC64;
```

Query Match   8.8%; Score 163; DB 1; Length 3210;
Best Local Similarity 21.8%; Pred. No. 3.9;

Matches   84;   Conservative   56;   Mismatches   147;   Indels   98;   Gaps   12;

```
Qy    44  KEVEALLSSIDEIAAKAIG-----KKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLK 100
              |:| |:|          | | :|         |  :|| |:| | |||||:|||:|
Db  2257  RELQMSEENQELVILDAENSKAEVETLKTQIEEMARSLKVFELDL----VTLRSEKENLT 2312

Qy   101  N--------------EGLKEKIDAAKKCSETFTNKLKEKHTDLGKEGVTDADAKEAIL 144
                          ||||||  |  | |  |:|| : |  |:  |
Db  2208  EEVHQLRGIEKLRVIEADEKKQLHIAEKLKERERENDSL-------KDKVENIE 2256

Qy   145  KANGTKTKGAEELGKLFESVEVL--------SKAAKEMLANSVKELTSPVVAES 190
              |||:|| ||  | |:| | |             |:| ||:||::|:|::|
Db  2313  KQIQEKQGQLSELDKLLSSFKSLLEEKEQAETQIKEESKTAVEMLQNQLKELNEAV----- 2368

Qy   191  PKKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLTSIDEL 250
              | |:|  |:|  |      | :|| :|| |:| | :| |  | | |::| ||:|:||
Db  2369  ----AALCGDQEIMKATEQSLDPPIEEHQLRNSIEKLRARLEADEKK 2412

Qy   251  AKAIGKKIKNDVSLDNEADHNGSLISGAYLISNL------ITKKISAIKDSGELKAEIE 303
              | | |:| | |  | ::|||||:|: |  |            |:: | | | |::|
Db  2413  QLCVLQQLK-------ESEHHADLLKGR--VENLERELEIARTNQEHAALEAENSKGEVE 2463

Qy   352  EKLFESVKNLSKAAKEMLTNSVKEL 376
              :| |:| | |::| : | :||:||
Db  2464  TLKAKIEGMTQSLRGLELDVVTIRSEKEDLTNELQKEQERISELEIINSSFENILQEKEQ 2523

Db  2524  EKV--QMKEKSSTAMEMLQTOLKEL 2546
```

RESULT 15
```
ID   MYSB_CAEEL     STANDARD;     PRT;     1966 AA.
AC   P02566;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   MYOSIN HEAVY CHAIN B (MHC B).
GN   UNC-54 OR MYO-4.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=83273600; PubMed=6576334;
RA   Karn J., Brenner S., Barnett L.;
RT   "Protein structural domains in the Caenorhabditis elegans unc-54
RT   myosin heavy chain gene are not separated by introns.";
RL   Proc. Natl. Acad. Sci. U.S.A. 80:4253-4257(1983).
RN   [2]
RP   SEQUENCE OF 850-1966 FROM N.A.
RX   MEDLINE=82272395; PubMed=7202124;
RA   McLachlan A.D., Karn J.;
RT   "Periodic charge distributions in the myosin rod amino acid sequence
RT   match cross-bridge spacings in muscle.";
RL   Nature 299:226-231(1982).
RN   [3]
RP   SEQUENCE OF 1876-1966 FROM N.A.
RX   MEDLINE=83228092; PubMed=6571695;
RA   Wills N., Gesteland R.F., Karn J., Bolten S.,
RA   Waterston R.H.;
RT   "The genes sup-7 X and sup-5 III of C. elegans suppress amber
RT   nonsense mutations via altered transfer RNA.";
RL   Cell 33:575-583(1983).
CC   -!- FUNCTION: MUSCLE CONTRACTION.
CC   -!- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
CC       HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
CC       AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
CC   -!- SUBCELLULAR LOCATION: THICK FILAMENTS OF THE MYOFIBRILS.
CC   -!- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CC       CYCLES OF A 28-RESIDUE TAIL REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
```

```
CC       CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC   -!- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
CC       ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
CC   -!- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
CC       MEROMYOSIN (LMM) AND 1 HEAVY MEROMYOSIN (HMM). IT CAN LATER BE
CC       SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
CC       SUBFRAGMENT (S2).
CC   -!- MISCELLANEOUS: THERE ARE FOUR DIFFERENT MYOSIN HEAVY CHAINS IN C.
CC       ELEGANS.
CC   -!- MISCELLANEOUS: MHC A AND MHC B ARE FOUND EXCLUSIVELY IN THE BODY
CC       WALL MUSCLE. THEY CO-ASSEMBLE INTO BODY WALL THICK FILAMENT.
CC   -!- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC   --------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------
DR   EMBL; J01050; AAA28124.1; -.
DR   EMBL; V01494; CAA24738.1; -.
DR   PIR; A02992; MKKW.
DR   HSSP; P08799; 1MND.
DR   InterPro; IPR002928; Myosin_tail.
DR   InterPro; IPR001609; myosin_head.
DR   Pfam; PF00063; Myosin_head; 1.
DR   Pfam; PF01576; Myosin_tail; 1.
DR   PRINTS; PR00193; MYOSINHEAVY.
DR   ProDom; PD000355; Myosin_head; 1.
DR   SMART; SM00242; MYSc; 1.
KW   Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
KW   ATP-binding; Methylation; Alkylation; Multigene family.
FT   DOMAIN         1    850       MYOSIN HEAD-LIKE.
FT   DOMAIN       851   1966       COILED COIL (POTENTIAL).
FT   DOMAIN       851   1164       ALPHA-HELICAL TAILPIECE (S2).
FT   DOMAIN      1165   1176       HINGE.
FT   DOMAIN      1165   1966       LIGHT MEROMYOSIN (LMM).
FT   NP_BIND      177    184       ATP (BY SIMILARITY).
FT   DOMAIN       665    687       ACTIN-BINDING.
FT   DOMAIN       769    783       ACTIN-BINDING.
FT   MOD_RES      128    128       METHYLATION (TRI-) (POTENTIAL).
FT   MOD_RES      705    705       ALKYLATION (SH-1).
FT   MOD_RES      705    705       ALKYLATION (SH-2).
FT   MOD_RES      715    715       METHYLATION (SH-2).
FT   CONFLICT    1337   1337       E -> R (IN REF. 2).
FT   CONFLICT    1880   1880       L -> L (IN REF. 2).
SQ   SEQUENCE   1966 AA;  225125 MW;  B66F0B82FE27B67F CRC64;
```

```
Query Match             8.5%;  Score 157.5;  DB 1;  Length 1966;
Best Local Similarity   23.0%;           Pred. No. 4;
Matches   90;  Conservative   69;  Mismatches  147;  Indels   85;  Gaps   18;

Qy    16  SADESVKGPNLTEINKKITDSNAVLLAVK---EVEALLSSIDEI--AAKALKIHQ-NN 69
              ::|| | |  :| :    : | || | |         |:||||||:|    ||:||:| ||
Db  1415  ALDAANSKNASLEKTKSRLVGDLDDAQVDVERANGVASALEKKQKGFDKIIDEWRKKTDD 1474

Qy    70  GLDTENNHNGSL-LAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNKLKEKHTD 128
              |||| |||:|| | | || |||:|||:|   ||:|  |  |: | |  || |:|:|:||
Db  1358  SLEEEIEGKN--EILRQLSKANADIQQWKARFEGEGLLKA-DELEDAKRRQAQKINELQE 1414

Qy   129  LGKEGVTDADAKEAILKANGI---KTKGAEELGKLFESVEVLSKAAKEMLANSVKELTSP 185
              ||||| |||||||:| ||:        || | | |||| || | | |::||||:||||
Db  1475  LAAE---LDGAQRDLRNTSTDLFRAKNAQE-ELAEVVEGLRRENKS-LSQEIKDLTDQ 1527

Qy   186  VVAESPKKPSMVNNSGKDGNTSANSADESVKGPNLTEISKK-----------ITESNA 232
              ||| |  |  |  ||:|   |||    |:     |   ||
Db  1528  L---------GEGG----RSVHEMQKIIRRLEIEKEELQHALDEAEAALEAEESK 1569

Qy   233  VVLAVKEVETLLTSIDELAKAIGKKIKN------------DVSLDNEADHNGSLISGAYL 280
              ||| |||||:||||| ||| | |||||:                ||:||||||:| ||
Db  1570  VLRAQVEVSQIRSEIEKRIQEKEEEFENTRKNHARALESMQOASLETEAKGKAELLR---- 1625
```

```
Qy    281  ISNLITKKISAIKDSGELKAEIEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKT  340
                 | ||:    |  |||:   :| :: ||:  |: :|     ||
Db   1626  ----IKKKLEG--DINELEIALDHANKANADAQKNLKR----------YOEQVRELQLQV  1669

Qy    341  NNDKTKGADELEKLFESVK--NLSKAAKEML  369
                 :: ||||    ::|: ||:|    |: :|| |
Db   1670  EEEQRNGADTREQFFNAEKRATLLQSEKEEL  1700
```

Search completed: March 18, 2002, 10:11:45
Job time: 963 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 09:56:46 ; Search time 68.77 Seconds
                                    (without alignments)
                                      418.700 Million cell updates/sec

Title: US-09-596-746A-28
Perfect score: 1859
Sequence: 1 MACNNSGKDGNTSANSADES.........KNLSKAAKEMLTNSVKELTS 378

Scoring table: BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   PIR_68:*
                 1:   pir1:*
                 2:   pir2:*
                 3:   pir3:*
                 4:   pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	% Query Match	Length	DB	ID	Description
1	945	50.8	210	2	G70218	outer surface prot
2	938	50.5	210	2	S69927	outer surface prot
3	879	47.3	210	2	I40144	outer surface prot
4	853	45.9	177	2	I40129	outer surface prot
5	774.5	41.7	211	2	I40145	outer surface prot
6	773.5	41.6	193	2	S70279	outer surface prot
7	700.5	37.7	193	2	S70287	outer surface prot
8	695.5	37.4	211	2	S69918	outer surface prot
9	692.5	37.3	211	2	I40277	outer surface prot
10	690.5	37.1	193	2	S70280	outer surface prot
11	686.5	36.9	211	2	I40268	outer surface prot
12	686	36.9	190	2	S70273	outer surface prot
13	685.5	36.9	209	2	I40273	outer surface prot
14	683.5	36.8	191	2	S70278	outer surface prot
15	683.5	36.8	191	2	I40153	outer surface prot
16	675	36.3	194	2	S70277	outer surface prot
17	674.5	36.3	209	2	S69917	outer surface prot
18	670	36.0	212	2	S70254	outer surface prot
19	666.5	35.9	193	2	S70276	outer surface prot
20	666	35.8	193	2	S70281	outer surface prot
21	664.5	35.7	211	2	S70274	outer surface prot
22	664	35.7	211	2	I40278	outer surface prot
23	662.5	35.6	212	2	S69922	outer surface prot
24	662.5	35.6	191	2	S70284	outer surface prot
25	662	35.6	193	2	S70265	outer surface prot
26	662	35.6	194	2	S70268	outer surface prot
27	661.5	35.6	211	2	S69932	outer surface prot
28	657.5	35.4	209	2	I40142	outer surface prot
29	655	35.2	212	2	I40279	outer surface prot
30	650	35.0	212	2	I40143	outer surface prot
30	646	34.7	214	2	S69916	outer surface prot
31	641.5	34.5	193	2	S70286	outer surface prot
32	638	34.3	194	2	S70289	outer surface prot
33	635.5	34.2	191	2	S70288	outer surface prot
34	635.5	34.2	203	2	I40108	outer surface prot
35	631	33.9	212	2	S20543	outer surface prot
36	630.5	33.9	209	2	I40281	outer surface prot
37	630.5	33.9	211	2	S70285	outer surface prot
38	630	33.9	192	2	S69925	outer surface prot
39	628	33.8	210	2	S69923	outer surface prot
40	627	33.7	212	2	S69921	outer surface prot
41	627	33.7	210	2	S69926	outer surface prot
42	622.5	33.5	209	2	S69920	outer surface prot
43	622	33.5	210	2	I40272	outer surface prot
44	619	33.3	210	2	S70259	outer surface prot
45	617.5	33.2	193	2		

ALIGNMENTS

RESULT   1
G70218
outer surface protein C - Lyme disease spirochete
outer surface protein C - Lyme disease spirochete (Lyme disease spirochete)
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 21-Jul-2000
C;Accession: G70218; I40269; S37726; S70281
R;Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; Wh
son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vu
; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997
A;Authors: Smith, H.O.; Venter, J.C.
A;Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.
R;Reference number: A70100; MUID:98065943
A;Accession: G70218
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-210 <KLE>
A;Cross-references: GB:AED00792; NID:g3253098; PIDN:AAC66329.1; PID:g2689901; TIGR:BB
A;Experimental source: strain B31
R;Fukunaga, M.; Hamase, A.
J. Clin. Microbiol. 33, 2415-2420, 1995
A;Title: Outer surface protein C gene sequence analysis of Borrelia burgdorferi sensu
A;Reference number: I40269; MUID:96025162
A;Accession: I40269
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-210 <RES>
A;Cross-references: GB:D49497; NID:g707092; PIDN:BAA08457.1; PID:g769684
R;Jauris-Heipke, S.; Fuchs, R.; Motz, M.; Preac-Mursic, V.; Schwab, E.; Soutschek, E.
Med. Microbiol. Immunol. 182, 37-50, 1993
A;Title: Genetic heterogenity of the genes coding for the outer surface protein C (Os
A;Reference number: S37726; MUID:93268136
A;Accession: S37726
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-210 <JAU>
A;Cross-references: EMBL:X69596; NID:g311391; PIDN:CAA49306.1; PID:g311392
R;Livey, I.; Gibbs, C.P.; Schuster, R.; Dorner, F.
Mol. Microbiol. 18, 257-269, 1995
A;Title: Evidence for lateral transfer and recombination in OspC variation in Lyme di
A;Reference number: S70255; MUID:96296448
A;Accession: S70281
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 19-210 <LIV>
A;Cross-references: EMBL:L42887; NID:g858715; PIDN:AAB36995.1; PID:g1695212
C;Genetics:
A;Experimental source: strain Ip2
C;Genetics:
A;Gene: ospC
C;Superfamily: Lyme disease spirochete surface protein C

RESULT 3
I40144

Query Match   Best Local Similarity   50.8%;   Score 945; DB 2; Length 210;
                                97.9%; Pred. No. 4e-37;
Matches 190; Conservative   3; Mismatches   1; Indels   0; Gaps   0;

Qy   1 MACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKA 60
          ::||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   17 ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA 60

Qy   61 IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   77 IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN 136

Qy   121 KLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   137 KLKEKHTDLGKEGVTDADAKEAILKINGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 196

Qy   181 ELTSPVVAESPKKP 194
          |||||||||||
Db   197 ELTSPVVAESPKKP 210

● ospc
Superfamily: Lyme disease spirochete surface protein C

RESULT 2
S69927

outer surface protein C precursor - Lyme disease spirochete (strain Pka)
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
A;Variety: strain Pka
C;Date: 15-Feb-1997 #sequence_revision 27-Feb-1997 #text_change 26-May-2000
R;Jauris-Heipke, S.; Liegl, G.; Preac-Mursic, V.; Roessler, D.; Schwab, E.; Soutschek, E
J;Clin. Microbiol. 33, 1860-1866, 1995
A;Title: Molecular analysis of genes encoding outer surface protein C (OspC) of Borrelia
A;Reference number: I40047; MUID:95395018
A;Accession: S69927
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-210 <JAU>
A;Cross-references: EMBL:X69589
A;Experimental source: strain Pka
R;Jauris, S.
submitted to the EMBL Data Library, February 1994
A;Reference number: S72669
A;Accession: S72669
A;Molecule type: DNA
A;Residues: 1-124,'D',126-139,'E',141-210 <JAW>
A;Cross-references: EMBL:X69589
C;Genetics:
Gene: ospc
Superfamily: Lyme disease spirochete surface protein C

Query Match   Best Local Similarity   96.9%;   Score 938; DB 2; Length 210;
                                96.9%; Pred. No. 8.5e-37;
Matches 188; Conservative   5; Mismatches   1; Indels   0; Gaps   0;

Qy   1 MACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKA 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA 60

Qy   61 IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIIKQKLDGLKNEGLKEKIEAAKKCSETFTN 120

Qy   121 KLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   137 KLKDKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 180

Qy   181 ELTSPVVAESPKKP 194
          ||||||||||||||
Db   197 ELTSPVVAESPKKP 196

---

outer surface protein C precursor - Lyme disease spirochete
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 21-Jul-2000
R;Stevenson, B.; Barthold, S.W.
FEMS Microbiol. Lett. 124, 367-372, 1994
A;Title: Expression and sequence of outer surface protein C among North American isol
A;Reference number: I40143; MUID:95154673
A;Accession: I40144
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-210 <RES>
A;Cross-references: EMBL:U04281; NID:g434663; PIDN:AAB37001.1; PID:g1695218
R;Livey, I.; Gibbs, C.P.; Schuster, R.; Dorner, F.
Mol. Microbiol. 18, 257-269, 1995
A;Title: Evidence for lateral transfer and recombination in OspC variation in Lyme di
A;Reference number: S70255; MUID:96296448
A;Accession: S70282
A;Molecule type: DNA
A;Residues: 19-210 <LIV>
A;Cross-references: EMBL:L42893; NID:g858721; PIDN:AAC43297.1; PID:g434664
A;Experimental source: strain 297
A;Gene: ospc
C;Superfamily: Lyme disease spirochete surface protein C

Query Match   Best Local Similarity   47.3%;   Score 879; DB 2; Length 210;
                                99.4%; Pred. No. 4.3e-34;
Matches 180; Conservative   0; Mismatches   1; Indels   0; Gaps   0;

Qy   198 NNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLTSIDELAKAIGKK 257
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   20 NNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLTSIDELAKAIGKK 79

Qy   258 IKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAEIEKAKKCSEEFTAKLK 317
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   80 IKNDVSLDNEADHNGSLISGAYLISTLITKKISAIKDSGELKAEIERAKKCSEEFTAKLK 139

Qy   318 GEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLSKAAKEMLTNSVKELT 377
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   140 GEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLSKAAKEMLTNSVKELT 199

Qy   378 S 378
          |
Db   200 S 200

RESULT 4
I40129

outer surface protein C - Lyme disease spirochete (strain TXGW) (fragment)
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 26-May-2000
R;Theisen, M.; Borre, M.; Mathiesen, M.J.; Mikkelsen, B.; Lebech, A.M.; Hansen, K.
J;Bacteriol. 177, 3036-3044, 1995
A;Title: Evolution of the Borrelia burgdorferi outer surface protein OspC.
A;Reference number: I40104; MUID:95286481
A;Accession: I40129
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-177 <RES>
A;Cross-references: EMBL:X84783; NID:g793825; PIDN:CAA59254.1; PID:g793826
A;Gene: ospc
C;Superfamily: Lyme disease spirochete surface protein C

Query Match   Best Local Similarity   45.9%;   Score 853; DB 2; Length 177;
                                98.3%; Pred. No. 5.5e-33;
Matches 174; Conservative   1; Mismatches   2; Indels   0; Gaps   0;

```
Qy   11  NTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGKKIHQNNG 70
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  NTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGKKIHQNNG 60

Qy   71  LDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSEFTFNKLKEKHTDLG 130
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  LDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSEFTFNKLKEKHTDLG 120

Qy  131  KEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKELTSPVV 187
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121  KRGVTDADAKEAILKNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKELTSPVV 177
```

```
RESULT  5
I40145
outer surface protein C precursor - Lyme disease spirochete
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 21-Jul-2000
C;Accession: I40145
R;Stevenson, B.; Barthold, S.W.
FEMS Microbiol. Lett. 124, 367-372, 1994
Title: Expression and sequence of outer surface protein C among North American isolate
Reference number: I40143; MUID:95154673
A;Accession: I40145
A;Status: preliminary: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-211 <RES>
A;Cross-references: EMBL:U04282; NID:g2314881; PIDN:AAC45540.1; PID:g434666
C;Genetics:
A;Gene: ospC
C;Superfamily: Lyme disease spirochete surface protein C
```

```
Query Match             41.7%;   Score 774.5;   DB 2;   Length 211;
Best Local Similarity   81.0%;   Pred. No. 2.7e-29;
Matches 158;  Conservative  14;  Mismatches  22;  Indels  1;  Gaps  1;
```

```
Qy    1  MACNNSGKDGN-TSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAK 59
         :||||||||| |||||||||||||||||||||||||||||||||||||||:|||:|||
Db   17  ISCNNSGKDGNAASTNPADESVKGPNLTEISKKITDSNAVLLAVKEVEGALLTSIDELATK 76

Qy   60  AIGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSEFT 119
         |||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db   77  AIGKKIHQNNGLDTENNHNGSLLAGAYAISTLITQKLGGLKNEELKEKIAAVKKCSEEFT 136

Qy  120  NKLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSV 179
         ||||:||::|||:|:| :|| :||| :||||||||||||||||||||||||||||||||
Db  137  NKLKSSHTELGKQDAQDDDAKKAILRPHNTKDKGAEELDKLFKAVENLSKAAKEMLSNSV 196

Qy  180  KELTSPVVAESPKKP 194
         ||||||||||||||
Db  197  KELTSPVVAESPKKP 211
```

```
RESULT  6
S70279
outer surface protein C - Lyme disease spirochete
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 26-May-2000
C;Accession: S70279
R;Livey, I.; Gibbs, C.P.; Schuster, R.; Dorner, F.
Mol. Microbiol. 18, 257-269, 1995
A;Title: Evidence for lateral transfer and recombination in OspC variation in Lyme disea
A;Reference number: S70255; MUID:96296448
A;Accession: S70279
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-193 <LIV>
A;Cross-references: EMBL:L42898; NID:g858729; PIDN:AAB37007.1; PID:g1695223
C;Genetics:
A;Gene: ospC
C;Superfamily: Lyme disease spirochete surface protein C
```

```
RESULT  7
S70287
outer surface protein C - Lyme disease spirochete
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 26-May-2000
C;Accession: S70287
R;Livey, I.; Gibbs, C.P.; Schuster, R.; Dorner, F.
Mol. Microbiol. 18, 257-269, 1995
A;Title: Evidence for lateral transfer and recombination in OspC variation in Lyme di
A;Reference number: S70255; MUID:96296448
A;Accession: S70287
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-193 <LIV>
A;Cross-references: EMBL:L42895; NID:g858723; PIDN:AAB37003.1; PID:g1695220
C;Genetics:
A;Gene: ospC
C;Superfamily: Lyme disease spirochete surface protein C
```

```
Query Match             41.6%;   Score 773.5;   DB 2;   Length 193;
Best Local Similarity   81.9%;   Pred. No. 2.7e-29;
Matches 158;  Conservative  11;  Mismatches  23;  Indels  1;  Gaps  1;
```

```
Qy    3  CNNSGKDGN-TSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKAI 61
         |||||||| |||||||||||||||||||||||||||||||||||||||||:|||:|||
Db    1  CNNSGKDGNAASTNPADESVKGPNLTEISKKITDSNTVVLAVKEVEGALLTSIDELATK 60

Qy   62  GKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNK 121
         ||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db   61  GKKIHQNNGLDTENNHNGSLLAGAYAISTLITQKLGGLKNEELKEKIAAVKKCSEEFTNK 120

Qy  122  LKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 181
         |:||::|||:: |:| :|| :||| :|||||||||||||||||||||||||||:|||
Db  121  LKSSHTELGKQDAQDDDAKKAILRPHNTKDKGAEELDKLFKPVENLSKAAKEMLSNSVKE 180

Qy  182  LTSPVVAESPKKP 194
         |||||||||||||
Db  181  LTSPVVAESPKKP 193
```

```
RESULT  8
S69918
outer surface protein C precursor - Lyme disease spirochete (strain PBre)
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
A;Variety: strain PBre
C;Date: 06-Dec-1996 #sequence_revision 14-Feb-1997 #text_change 26-May-2000
C;Accession: S69918; S72674; I40103
R;Jauris-Heipke, S.; Liegl, G.; Preac-Mursic, V.; Roessler, D.; Schwab, E.; Soutschek
```

```
Query Match             37.7%;   Score 700.5;   DB 2;   Length 193;
Best Local Similarity   75.1%;   Pred. No. 5.9e-26;
Matches 145;  Conservative  17;  Mismatches  30;  Indels  1;  Gaps  1;
```

```
Qy    3  CNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKAIG 62
         |||||||||||||||||||||||||||||||||||| |||||||||:|||:|||:||
Db    1  CNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASTDELATKAIG 60

Qy   63  KKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKN-EGLKEKIDAAKKCSETFTNK 120
         |||||||||||||||| ||||:|||:||||| |||||| |||||| |||||||||| |
Db   61  KKIQONGGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKEKIENAKKCSEDFTKK 120

Qy  122  LKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 181
         |:| ::|||::|::| |  |  |:|| :|:| ||||  ||||||:||||||||||||||
Db  121  LEGEHAQLGIENVDENAKKAILITDAAKDKGAAELEKLFKAVENLAKAAKEMLANSVKE 180

Qy  182  LTSPIVAESPKNP 194
         |||:||||||| |
Db  181  LTSPIVAESPKNP 193
```

J. Clin. Microbiol. 33, 1860-1866, 1995
A;Title: Molecular analysis of genes encoding outer surface protein C (OspC) of Borrelia
A;Reference number: I40047; MUID:95395018
A;Accession: S69918
A;Molecule type: DNA
A;Residues: nucleic acid sequence not shown
A;Status: nucleic acid sequence not shown
A;Residues: 1-152, 'E', 154-211 <JAU>
A;Cross-references: 1-211 <JAU>
A;Molecule type: DNA
A;Cross-references: EMBL:X81522; NID:g872021; PIDN:CAA57242.1; PID:g872022
A;Experimental source: strain PBre
R;Roessler, D.
submitted to the EMBL Data Library, September 1994
A;Reference number: S72674
A;Accession: S72674

ne: ospc
perfamily: Lyme disease spirochete surface protein C

RESULT 9
Query Match
Best Local Similarity 37.4%; Score 695.5; DB 2; Length 211;
Matches 152; Conservative 12; Mismatches 29; Indels 3; Gaps 3;

Qy   1 MACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKA 60
         ::|||||| |:: |||||||||||| |||||||||||||||||||:||::|||| |||
Db   17 ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEL-AKA 75

Qy   61 IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLK -NEGLKEKIDAAAKKCSETFT 119
         |||| ||:| |||||| |||||||||||||||||||||| | ||||||||||||:|
Db   76 IGKKIKNDGSLGDEANHNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEEFS 135

Qy 120 NKLKEKHTDLGKEGVTDADAKEAILKANGT~KTKGAEELGKLFESVEVLSKAAKEMLANS 178
         | | : ||| || ||| | ||||||| |||||||||||||||||||||||||||||
Db 136 TKLKDNHAQLGIQGVTDDNAKKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANSVK 195

Qy 179 VKELTSPVVAESPKKP 194
         |||||||||||||||
Db 196 VKELTSPVVAESPKKP 211

RESULT 10
S70280
outer surface protein C - Lyme disease spirochete
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 26-May-2000
C;Accession: S70280
C;Genetics:
A;Experimental source: strain ZS7
A;Cross-references: EMBL:L42868; NID:g858735; PIDN:g1695226
A;Genetics:
A;Gene: ospC
C;Superfamily: Lyme disease spirochete surface protein C

Query Match
Best Local Similarity 37.1%; Score 690.5; DB 2; Length 193;
Matches 151; Conservative 10; Mismatches 30; Indels 3; Gaps 3;

Qy   3 CNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKAIG 62
         |||||||||||||| |:: |||||||| |||||||||||||||:||::|||| ||| |
Db   1 CNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEL-AKAIG 59

Qy   63 KKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLK -NEGLKEKIDAAAKKCSETFTNK 121
         ||| ||:| |||||| |||||||||||||||||||| | ||||||||||||:| |
Db   60 KKIKNDGSLGDEANHNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEEFSTK 119

Qy 122 LKEKHTDLGKEGVTDADAKEAILKANGT~KTKGAEELGKLFESVEVLSKAAKEMLANSVK 180
         | : ||| || ||| | ||||||| |||||||||||||||||||||||||||||
Db 120 LKDNHAQLGIQGVTDENAKKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANSVK 179

Qy 181 ELTSPVVAESPKKP 194
         |||||||||||||
Db 180 ELTSPVVAESPKKP 193

RESULT 11
I40268
outer surface protein C precursor - Lyme disease spirochete
A;Title: Homology between Borrelia burgdorferi OspC and members of the family of Borr
A;Reference number: I40268; MUID:94259285
A;Accession: I40268
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-211 <RES>
A;Cross-references: GB:L25413; NID:g495735; PIDN:AAA22956.1; PID:g495736
A;Genetics:
A;Gene: ospC
C;Superfamily: Lyme disease spirochete surface protein C

Query Match
Best Local Similarity 36.9%; Score 686.5; DB 2; Length 211;
Matches 144; Conservative 18; Mismatches 32; Indels 1; Gaps 1;

RESULT 9 (continued below)
I40277
outer surface protein C precursor - Borrelia afzelii
C;Species: Borrelia afzelii
te: 04-Sep-1997 #sequence_revision 04-Sep-1997 #text_change 26-May-2000
cession: I40277
R;Fukunaga, M.; Hamase, A.
J. Clin. Microbiol. 33, 2415-2420, 1995
A;Title: Outer surface protein C gene sequence analysis of Borrelia burgdorferi sensu la
A;Reference number: I40277
A;Accession: I40277
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-211 <RES>
A;Cross-references: GB:D49501; NID:g707096; PIDN:BAA08461.1; PID:g769688
C;Superfamily: Lyme disease spirochete surface protein C

Query Match
Best Local Similarity 37.8%; Score 692.5; DB 2; Length 211;
Matches 146; Conservative 20; Mismatches 27; Indels 3; Gaps 3;

Qy   1 MACNNSGKDGNT-SANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAK 59
         ::|||||| |:: |||||||||||||||||||:|:||||||||||||||||||:||::|
Db   17 ISCNNSGKGGDSASTNPTDESAKGPNLTEISKKITDSNAFVLAVKEVETLVSSIDELATK 76

Qy   60 AIGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKN -EGLKEKIDAAAKKCSETF 118
         ||||| ||:| |||||| |||||||||||||||||||| || |||||||||||:|
Db   77 AIGKI-QNNGLAAEADKNGSLLAGAYAISTLIKQKLDGLKGLEGLNKEIAEAAKKCSEAF 135

Qy 119 TNKLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANS 178
         | ||::| ||| || ||| | ||||||||||||||||||||||||||||||||||

```
Qy    1 MACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKA  60
        ::||||||||||||||||||||||||||:||||||||||||||||||||:|:|:||:|| ||
Db   17 ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEVLLSSIDELAKKA  76

Qy   61 IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEG-LKEKIDAAKKCSEFFT  119
        ||||| |||:||:|||| |||||||||||||||||||||:|:|| ||:|:||||||||||||
Db   77 IGKKIDQNNALGTLDNNNGSLLAGAYAISALITEKLSSIKDSGELKAEIEKAKKCSESFT  136

Qy  120 NKLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSV  179
        :||:||:|||:|::||:|||:|||||:|:|::||:||||:|:|:|:|:||||:||:|||||
Db  137 KKLSDNQAELGTENATDDNAKKAILKTHNAKDKGAEELVKLSESVAGLLKAAQAILANSV  196

Qy  180 KELTSPVVAESPKKP  194
        |||||||||||||||
Db  197 KELTSPVVAESPKKP  211
```

RESULT 12
S70273
outer surface protein C - Lyme disease spirochete
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
...ate: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 26-May-2000
C;Accession: S70273
A;Molecule type: DNA
A;Status: nucleic acid sequence not shown
A;Residues: 1-190 <LIV>
R;Livey, I.; Gibbs, C.P.; Schuster, R.; Dorner, F.
Mol. Microbiol. 18, 257-269, 1995
A;Title: Evidence for lateral transfer and recombination in OspC variation in Lyme disea
A;Reference number: S70255; MUID:96296448
A;Accession: S70273
A;Experimental source: strain VSDA
C;Genetics:
A;Gene: ospC
C;Superfamily: Lyme disease spirochete surface protein C

Query Match   36.9%; Score 686; DB 2; Length 190;
Best Local Similarity 75.6%; Pred. No. 2.7e-25;
Matches 146; Conservative 19; Mismatches 24; Indels 4; Gaps 3;

```
Qy    3 CNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKAIG  62
        |||| |||||||||||||||||:|||||||||||||||||:||||||||:|||:|||||| ||
Db    1 CNNSG--GDTASTNPDESAKGPDLTVISKKITDSNAFVLAVKEVEALLSSVDEL-AKAIG  57

Qy   63 KKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLIK-NEGLKEKIDAAKKCSETFTNK  121
        ||| :|||||::|||| :|||||||||||||:|||||||| |||:|:||||||||||||||
Db   58 KKIHQNNGLDTLSNQNGSLLAGAYAISTLITRKLDGLKGSEGLKAEIEAAKKCSEDFTKK  117

Qy  122 LKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE  181
        |||| |||:|::|||:||||:|:|:| ::|:||:||:||:|:|:|:|| |||||||||||||
Db  118 LKEKHTELGVAAATDDNAKKAILKANGDKTLGVEELEKLFKSVEKLSKAAQEALANSVQE  177

Qy  182 LTSPVVAESPKKP  194
        |||||||||||||
Db  178 LTSPVVAESPKKP  190
```

RESULT 13
I40273
outer surface protein C precursor - Borrelia afzelii
C;Species: Borrelia afzelii
C;Date: 04-Sep-1997 #sequence_revision 04-Sep-1997 #text_change 26-May-2000
C;Accession: I40273
R;Fukunaga, M.; Hamase, A.
J. Clin. Microbiol. 33, 2415-2420, 1995
A;Title: Outer surface protein C gene sequence analysis of Borrelia burgdorferi sensu la
A;Reference number: I40273
A;Accession: I40273
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA

A;Residues: 1-209 <RES>
A;Cross-references: GB:D49379; NID:g1041107; PIDN:BAA08377.1; PID:g1041108
C;Superfamily: Lyme disease spirochete surface protein C

Query Match   36.9%; Score 685.5; DB 2; Length 209;
Best Local Similarity 74.4%; Pred. No. 3.2e-25;
Matches 145; Conservative 21; Mismatches 26; Indels 3; Gaps 3;

```
Qy    1 MACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKA  60
        :|||||||:||:|||||||||||||||||:||||||||||:||||||||||||||:|||:| ||
Db   17 ISCNNSGKGGDIASTNPDESAKGPNLTEISKKITDSNAVVLAVKEVEALLSSIDEL-AKT  75

Qy   61 IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKN-EGLKEKIDAAKKCSETFT  119
        |||||:||:|::|| :||||||||||||||||||||||||| |||:|:|||||||||:|||
Db   76 IGKKI-EANGLGNEADKNGSLLAGAYAISTLIKQKLDGLKGLEGLNKEIAEAKKCSEAFT  134

Qy  120 NKLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSV  179
        :||:||:|||:|:::||:|:|||:||:|:|:||||::||:||:|:||:|||||||:||:||
Db  135 KKLQDSNADLGKHNATDADSKEAILKTNGTKTKGAKELEELFKSVESLSKAAKEALSNSV  194

Qy  180 KELTSPVVAESPKKP  194
        |||||||||||||||
Db  195 KELTSPVVAESPKKP  209
```

RESULT 14
S70278
outer surface protein C - Lyme disease spirochete
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 26-May-2000
C;Accession: S70278
R;Livey, I.; Gibbs, C.P.; Schuster, R.; Dorner, F.
Mol. Microbiol. 18, 257-269, 1995
A;Title: Evidence for lateral transfer and recombination in OspC variation in Lyme di
A;Reference number: S70255; MUID:96296448
A;Accession: S70278
A;Molecule type: DNA
A;Status: nucleic acid sequence not shown
A;Residues: 1-191 <LIV>
A;Cross-references: EMBL:L42871; NID:g858738; PIDN:AAB37014.1; PID:g1695229
C;Genetics:
A;Gene: ospC
C;Experimental source: strain VS461
C;Superfamily: Lyme disease spirochete surface protein C

Query Match   36.8%; Score 683.5; DB 2; Length 191;
Best Local Similarity 75.1%; Pred. No. 3.5e-25;
Matches 145; Conservative 19; Mismatches 26; Indels 3; Gaps 3;

```
Qy    3 CNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKAIG  62
        |||| ::|:|:||:|||||||||||||||:||||||||||:||||||||||||||:|||| ||
Db    1 CNNSG-GGDIASTNPDESAKGPNLTEISKKITDSNAVVLAVKEVEALLSSIDEL-AKTIG  59

Qy   63 KKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKN-EGLKEKIDAAKKCSETFTNK  121
        |||:||:|::|| :||||||||||||||||||||||||| |||:|:|||||||||:|||||
Db   60 KKI-EANGLGNEADKNGSLLAGAYAISTLIKQKLDGLKGLEGLNKEIAEAKKCSEAFTKK  118

Qy  122 LKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE  181
        |||| |||:|::|||:|:||||:||:|:|:||||::||:||:|:||:|||||||:||:|||
Db  119 LQDSNADLGKHNATDADSKEAILKTNGTKTKGAKELEELFKSVESLSKAAKEALSNSVKE  178

Qy  182 LTSPVVAESPKKP  194
        |||||||||||||
Db  179 LTSPVVAESPKKP  191
```

RESULT 15
I40153
outer surface protein C - Lyme disease spirochete (fragment)
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 26-May-2000

C;Accession: I40153
R;Fung, B.P.; McHugh, G.L.; Leong, J.M.; Steere, A.C.
Infect. Immun. 62, 3213-3221, 1994
A;Title: Humoral immune response to outer surface protein C of Borrelia burgdorferi in L
A;Reference number: I40153; MUID:94314437
A;Accession: I40153
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-191 <RES>
A;Cross-references: EMBL:U08284; NID:g469561; PIDN:AAA21460.1; PID:g469562
C;Genetics:
A;Gene: ospC
C;Superfamily: Lyme disease spirochete surface protein C

Query Match   36.8%;   Score 683.5;   DB 2;   Length 191;
Best Local Similarity   74.9%;   Pred. No. 3.5e-25;
tches 143;   Conservative   17;   Mismatches   30;   Indels   1;   Gaps   1;

```
Qy      5 NSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGKK 64
            ||| ||||||||||||||||| :|||:||| || ||||||| ::|||||:| |||||
Db      1 NSGKGGNTSANSADESVKGPNLTEISKKIPESNAVVLAVKEIETLLASIDELATKAIGKK 60

Qy     65 IHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKN-EGIKEKIDAAKKCSETFTNKLK 123
            | || | ||| :|||| ::|||||| | ||:|| |||| | ||:
Db     61 IQQNGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKEKIENAKKCSEDFTKKLE 120

Qy    124 EKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKELT 183
            :| | |  ||| :||:||| : |  ||| ||| ||||||||||||||||||||||||
Db    121 GEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLAKAAKEMLANSVKELT 180

Qy    184 SPVVAESPKKP 194
            ||:|||||||
Db    181 SPIVAESPKKP 191
```

Search completed: March 18, 2002, 09:56:47
Job time: 370 sec

```
                                        GenCore version 4.5
                                        Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on:              March 18, 2002, 10:10:53 ; Search time 124.19 Seconds
                                                 (without alignments)
                                                 445.212 Million cell updates/sec

Title:        US-09-596-746A-28
Perfect score: 1859
Sequence:     1 MACNNSGKDGNTSANSADES.........KNLSKAAKEMLTNSVKELTS 378
Scoring table: BLOSUM62
              Gapop 10.0 , Gapext 0.5
Searched:     473505 seqs, 14627229 residues

Total number of hits satisfying chosen parameters:     473505
Maximum DB seq length: 200000000
Minimum DB seq length: 0

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :  SPTREMBL_17:*
            1:  sp_archea:*
            2:  sp_bacteria:*
            3:  sp_fungi:*
            4:  sp_human:*
            5:  sp_invertebrate:*
            6:  sp_mammal:*
            7:  sp_mhc:*
            8:  sp_organelle:*
            9:  sp_phage:*
            10: sp_plant:*
            11: sp_rodent:*
            12: sp_virus:*
            13: sp_vertebrate:*
            14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                              SUMMARIES

ult                 %      Query
NO.   Score    Match  Length  DB  ID       Description
----  -----    -----  ------  --  ------   -----------
1     943      50.7   192     2   Q9S3P3   borrelia bu
2     927      49.9   191     2   P70818   borrelia bu
3     921      49.5   192     2   Q9R7B1   borrelia bu
4     908      48.8   193     2   Q9RR53   borrelia bu
5     893      48.0   200     2   Q9RQR9   borrelia bu
6     885      47.6   182     2   Q9R7B2   borrelia bu
7     885      47.6   191     2   Q31120   borrelia bu
8     879      47.3   192     2   Q9S3P2   borrelia bu
9     879      47.3   210     2   Q44719   borrelia bu
10    853      45.9   177     2   Q44999   borrelia bu
11    843.5    45.4   178     2   Q9REH7   borrelia bu
12    829      44.6   175     2   Q9R7B4   borrelia bu
13    802      43.1   163     2   Q9R7A9   borrelia bu
14    786      42.3   165     2   Q9AGB1   borrelia bu
15    774.5    41.7   211     2   Q44720   borrelia bu
16    773.5    41.6   193     2   P94237   borrelia bu
17    773      41.6   159     2   Q9R7B0   borrelia bu
18    707      38.0   193     2   Q31115   borrelia bu
19    700.5    37.7   193     2   P94234   borrelia bu
20    696.5    37.5   211     2   Q44977   borrelia bu
21    692.5    37.3   211     2   Q49576   borrelia af
22    690.5    37.1   193     2   P94242   borrelia bu
23    686.5    36.9   211     2   Q44976   borrelia bu
24    686      36.9   190     2   P94244   borrelia bu
25    686      36.9   212     2   P94244   borrelia bu
26    685.5    36.9   209     2   Q9KIM5   borrelia af
27    683.5    36.8   191     2   Q44671   borrelia af
28    683.5    36.8   191     2   Q44726   borrelia bu
29    681      36.6   212     2   P94245   borrelia bu
30    675      36.3   194     2   P94247   borrelia bu
31    670.5    36.1   194     2   Q31122   borrelia bu
32    670      36.0   212     2   Q44727   borrelia bu
33    667      35.9   201     2   P96571   borrelia ja
34    666.5    35.9   193     2   P94231   borrelia bu
35    666.5    35.9   211     2   Q9Z6C7   borrelia va
36    666      35.8   201     2   P96573   borrelia bu
37    664.5    35.7   193     2   P94233   borrelia bu
38    664.5    35.7   202     2   P96508   borrelia af
39    664.5    35.7   209     2   Q44883   borrelia ja
40    664      35.7   211     2   Q49577   borrelia af
41    664      35.7   203     2   050624   borrelia af
42    664      35.7   212     2   Q44670   borrelia af
43    662.5    35.6   191     2   Q9S3P0   borrelia bu
44    662.5    35.6   193     2   P94226   borrelia bu
45    662      35.6   194     2   P94229   borrelia bu
```

                              ALIGNMENTS

```
RESULT 1

ID     Q9S3P3
AC     Q9S3P3;
DT     01-MAY-2000 (TrEMBLrel. 13, Created)
DT     01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT     01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE     OUTER SURFACE PROTEIN C (FRAGMENT).
OS     Borrelia burgdorferi (Lyme disease spirochete).
OC     Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX     NCBI_TaxID=139;
RN     [1]
RP     SEQUENCE FROM N.A.
RC     STRAIN=IP2;
RX     MEDLINE=96296448; PubMed=8709845;
RA     Livey I., Gibbs C.P., Schuster R., Dorner F.;
RT     "Evidence for lateral transfer and recombination in OspC variation in
RT     Lyme disease Borrelia.";
RL     Mol. Microbiol. 18:257-269(1995).
DR     EMBL; L42887; AAB36995.1; -.
DR     InterPro; IPR001800; Lipoprotein_6.
DR     Pfam; PF01441; Lipoprotein_6; 1.
DR     ProDom; PD001149; Lipoprotein_6; 1.
FT     NON_TER       1       1
FT     NON_TER     192     192
SQ     SEQUENCE    192 AA;    20287 MW;  11846F7AC84C7E3D CRC64;
```

```
Query Match             50.7%;    Score 943;   DB 2;  Length 192;
Best Local Similarity   99.0%;    Pred. No. 2.2e-36;
Matches  190;  Conservative  1;  Mismatches  1;  Indels  0;  Gaps  0;

Qy    3 CNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKAIG 62
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 CNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAIG 60

Qy   63 KKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAARKCSETFTNKL 122
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 KKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAARKCSETFTNKL 120

Qy  123 KEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKEL 182
```

```
Db  121 KEKHTDLGKEGVTDADAKEAILKINGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKEL 180
        ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Qy  183 TSPVVAESPKKP 194
        ||||||||||||
Db  181 TSPVVAESPKKP 192
```

RESULT 2

```
ID   P70818       PRELIMINARY;    PRT;   191 AA.
AC   P70818;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN (FRAGMENT).
GN   OSPC.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=2-1498 CA4;
RA   Probert W.S., Crawford M.R., Cadiz R.B., Lefebvre R.B.;
RL   Submitted (SEP-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; L81311; AAB06569.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
FT   NON_TER      1      1
SQ   SEQUENCE   191 AA;  20126 MW;  D2B9B1C82B4DC3C0 CRC64;

Query Match             49.9%;  Score 927;  DB 2;  Length 191;
Best Local Similarity   98.4%;  Pred. No. 1.2e-35;
  Matches  188;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy    4 NNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGK 63
        ||||  |||||||||||||||||||||||  ||||||||||||||||||||||||||||||
Db    1 NNSGKGGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGK 60

Qy   64 KIHQNNGLDTENNHINGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNKLK 123
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 KIHQNNGLDTENNHINGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNKLK 120

Qy  124 EKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKELT 183
        |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db  121 EKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKELT 180

Qy  184 SPVVAESPKKP 194
        |||||||||||
Db  181 SPVVAESPKKP 191
```

RESULT 3

```
ID   Q9R7B1       PRELIMINARY;    PRT;   192 AA.
AC   Q9R7B1;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C (FRAGMENT).
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TETS;
RX   MEDLINE=9747003; PubMed=9336916;
RA   Postic D., Foretz M., Baranton G.;
RT   "Borrelia burgdorferi sensu stricto, a bacterial species 'made in the
RT   U.S.A.'?";
RL   Int. J. Syst. Bacteriol. 47:1112-1117(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TETS;
RA   Marti-Ras N., Postic D., Foretz M., Baranton G.;
RL   Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U91798; AAB81895.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
FT   NON_TER    192    192
SQ   SEQUENCE   192 AA;  20297 MW;  677050?A20AAF764 CRC64;

Query Match             49.5%;  Score 921;  DB 2;  Length 192;
Best Local Similarity   97.9%;  Pred. No. 2.2e-35;
  Matches  186;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy    1 MACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKA 60
        ::||||||||||||||||||||||||||||| |::|||||||||||||||||||||||||
Db    3 ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA 62

Qy   61 IGKKIHQNNGLDTENNHINGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   63 IGKKIHQNNGLDTENNHINGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN 122

Qy  121 KLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 180
        |||||| ||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db  123 KLKEKHVDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 182

Qy  181 ELTSPVVAES 190
        ||||||||||
Db  183 ELTSPVVAES 192
```

RESULT 4

```
ID   Q9RR53       PRELIMINARY;    PRT;   193 AA.
AC   Q9RR53;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C (FRAGMENT).
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=OC1;
RX   MEDLINE=99091544; PubMed=9872945;
RA   Wang I.N., Dykhuizen D.E., Qiu W., Dunn J.J., Bosler E.M., Luft B.J.;
RT   "Genetic diversity of ospC in a local population of Borrelia
RT   burgdorferi sensu stricto.";
RL   Genetics 151:15-30(1999).
DR   EMBL; AF029860; AAB86543.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
FT   NON_TER    193    193
SQ   SEQUENCE   193 AA;  20502 MW;  5EFDB5AFF89865D1E CRC64;

Query Match             48.8%;  Score 908;  DB 2;  Length 193;
Best Local Similarity   97.9%;  Pred. No. 8.7e-35;
  Matches  183;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy    1 MACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKA 60
        ::||||||||||||||||||||||||||||| |::|||||||||||||||||||||||||
Db    7 ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA 66

Qy   61 IGKKIHQNNGLDTENNHINGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN 120
```

Db   67 IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFIN 126
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy   121 KLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   127 KLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 186

Qy   181 ELTSPVV 187
           |||||||
Db   187 ELTSPVV 193

RESULT 5
ID   Q9RQR9   PRELIMINARY;   PRT;   200 AA.
AC   Q9RQR9;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C PRECURSOR (FRAGMENT).
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
GN   OSPC.
RP   SEQUENCE FROM N.A.
RC   Plasmid cp26.
RC   STRAIN=2E7;
RX   MEDLINE=20002545; PubMed=10531219;
RA   Hofmeister E.K., Glass G.E., Childs J.E., Persing D.H.;
RT   "Population dynamics of a naturally occuring heterogeneous mixture of
RT   Borrelia burgdorferi clones.";
RL   Infect. Immun. 67:5709-5716(1999).
DR   EMBL; AF074464; AAD23911.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
KW   Plasmid.
FT   NON_TER   200   200
SQ   SEQUENCE   200 AA; 21307 MW; A22F158758BBBB6B CRC64;

Query Match   48.0%; Score 893; DB 2; Length 200;
Best Local Similarity   97.8%; Pred. No. 4.4e-34;
Matches 180; Conservative   3; Mismatches   1; Indels   0; Gaps   0;

Qy   1   MACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKA 60
           ::||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   17 ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA 76

Qy   61 IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFIN 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   77 IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFIN 136

Qy   121 KLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   137 KLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 196

Qy   181 ELTS 184
           ||||
Db   197 ELTS 200

RESULT 6
ID   Q9R7B2   PRELIMINARY;   PRT;   182 AA.
AC   Q9R7B2;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C (FRAGMENT).
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.

OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TETS;
RX   MEDLINE=97478003; PubMed=9336916;
RA   Ras N.M., Postic D., Foretz M., Baranton G.;
RT   "Borrelia burgdorferi sensu stricto, a bacterial species 'made in the
RT   U.S.A.'?";
RL   Int. J. Syst. Bacteriol. 47:1112-1117(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TETS;
RA   Marti-Ras N., Postic D., Foretz M., Baranton G.;
RL   Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U91797; AAB81894.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
FT   NON_TER   182   182
SQ   SEQUENCE   182 AA; 19202 MW; 422146F99AA57BF2 CRC64;

Query Match   47.6%; Score 885; DB 2; Length 182;
Best Local Similarity   98.9%; Pred. No. 9.1e-34;
Matches 180; Conservative   1; Mismatches   1; Indels   0; Gaps   0;

Qy   7   GKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGKKIH 66
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1   GKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGKKIH 60

Qy   67 QNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNKLKEKH 126
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 QNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNKLKEKH 120

Qy   127 TDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKELTSPV 186
           |||||||||||||||||||||:|||||||||||||||||||||||||||||||||||||
Db   121 TDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKELTSPV 180

Qy   187 VA 188
           ||
Db   181 VA 182

RESULT 7
ID   031120   PRELIMINARY;   PRT;   191 AA.
AC   031120;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C (FRAGMENT).
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=OC10;
RA   Wang I.-N., Dykhuizen D.E., Dunn J.J., Luft B.J.;
RL   Submitted (OCT-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF029869; AAB86552.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
FT   NON_TER   1   1
SQ   SEQUENCE   191 AA; 20502 MW; E4FB56BCB61740F9 CRC64;

Query Match   100.0%; Score 885; DB 2; Length 191;
Best Local Similarity 100.0%; Pred. No. 9.6e-34;
Matches 181; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy  198 NNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLTSIDELAKAIGKK 257
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   10 NNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLTSIDELAKAIGKK 69

Qy  258 IKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAEIEKAKKCSEEFTAKLK 317
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   70 IKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAEIEKAKKCSEEFTAKLK 129

Qy  318 GEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLSKAAKEMLINSVKELT 377
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  130 GEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLSKAAKEMLINSVKELT 189

Qy  378 S 378
        |
Db  190 S 190
```

RESULT 8
```
ID   Q9S3P2      PRELIMINARY;      PRT;   192 AA.
AC   Q9S3P2;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C (FRAGMENT).
GN   OSPC.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=297;
RX   MEDLINE=96299648; PubMed=8709845;
RA   Livey I., Gibbs C.P., Schuster R., Dorner F.;
RT   "Evidence for lateral transfer and recombination in OspC variation in
RT   Lyme disease Borrelia.";
RL   Mol. Microbiol. 18:257-269(1995).
DR   EMBL; L42893; AAB37001.1; -.
DR   Pfam; PF01441; Lipoprotein_6; -.
DR   ProDom; PD001149; Lipoprotein_6; 1.
DR   InterPro; IPR001800; Lipoprotein_6.
FT   NON_TER       1      1
FT   NON_TER     192    192
SQ   SEQUENCE   192 AA;  20472 MW;  46AC8F93E4DFED6C CRC64;

Query Match             47.3%;  Score 879;  DB 2;  Length 192;
Best Local Similarity   99.4%;  Pred. No. 1.8e-33;
Matches  180;  Conservative   0;  Mismatches   1;  Indels   0;  Gaps   0;
```

RESULT 9
```
ID   Q44999      PRELIMINARY;      PRT;   177 AA.
AC   Q44999;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C (FRAGMENT).
GN   OSPC.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TXGW;
RX   MEDLINE=95286481; PubMed=7768799;
RA   Theisen M., Borre M., Mathiesen M.J., Mikkelsen B., Lebech A.M.,
RA   Hansen K.;
RT   "Evolution of the Borrelia burgdorferi outer surface protein OspC.";
RL   J. Bacteriol. 177:3036-3044(1995).
DR   EMBL; X84783; CAA59254.1; -.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
DR   InterPro; IPR001800; Lipoprotein_6.
FT   NON_TER       1      1
FT   NON_TER     177    177
SQ   SEQUENCE   177 AA;  18800 MW;  15667189614E7A2D CRC64;

Query Match             45.9%;  Score 853;  DB 2;  Length 177;
```

```
Qy  198 NNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLTSIDELAKAIGKK 257
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   80 NNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLTSIDELAKAIGKK 257

Qy  258 IKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAEIEKAKKCSEEFTAKLK 317
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  198 IKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAEIEKAKKCSEEFTAKLK 317

Qy  318 GEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLSKAAKEMLINSVKELT 377
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  258 GEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLSKAAKEMLINSVKELT 377

Qy  378 S 378
        |
Db  378 S 378
```

RESULT 10
```
ID   Q44719      PRELIMINARY;      PRT;   210 AA.
AC   Q44719;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C.
GN   OSPC.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=HB19;
RX   MEDLINE=95154673; PubMed=7851744;
RA   Stevenson B., Barthold S.W.;
RT   "Expression and sequence of outer surface protein C among North
RT   American isolates of Borrelia burgdorferi.";
RL   FEMS Microbiol. Lett. 124:367-372(1994).
DR   EMBL; U04281; AAC43297.1; -.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
DR   InterPro; IPR001800; Lipoprotein_6.
SQ   SEQUENCE   210 AA;  22526 MW;  2D672991D584E4EE CRC64;

Query Match             47.3%;  Score 879;  DB 2;  Length 210;
Best Local Similarity   99.4%;  Pred. No. 2e-33;
Matches  180;  Conservative   0;  Mismatches   1;  Indels   0;  Gaps   0;
```

```
Qy  198 NNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLTSIDELAKAIGKK 257
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   20 NNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLTSIDELAKAIGKK 79

Qy  258 IKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAEIEKAKKCSEEFTAKLK 317
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   80 IKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAEIEKAKKCSEEFTAKLK 139

Qy  318 GEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLSKAAKEMLINSVKELT 377
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  140 GEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLSKAAKEMLINSVKELT 199

Qy  378 S 378
        |
Db  200 S 200
```

RESULT 11

Best Local Similarity 98.3%; Conservative 1; Pred. No. 2.5e-32; Mismatches 2; Indels 0; Gaps 0;
Matches 174;

```
Qy   11 NTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGKKIHQNNG  70
           |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db    1 NTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGKKIHQNNG  60

Qy   71 LDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKREKIDAAKKCSETFTNKLKEKHTDLG 130
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 LDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKREKIDAAKKCSETFTNKLKEKHTDLG 120

Qy  131 KEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKELTSPVV 187
           |||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db  121 KRGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKELTSPVV 177
```

ID   Q9REH7   PRELIMINARY;   PRT;   178 AA.
AC   Q9REH7;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C (FRAGMENT).
GN   OSPC.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94075528; PubMed=8253951;
RA   Theisen M., Frederiksen B., Lebech A.M., Vuust J., Hansen K.;
RT   "Polymorphism in ospC gene of Borrelia burgdorferi and
RT   immunoreactivity of OspC protein: implications for taxonomy and for
RT   use of OspC protein as a diagnostic antigen.";
RL   J. Clin. Microbiol. 31:2570-2576(1993).
DR   EMBL; X73622; CAA52001.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
FT   NON_TER   1   1
FT   NON_TER   178   178
SQ   SEQUENCE   178 AA; 18894 MW; D619AE4C646F14EB CRC64;

Query Match 45.4%; Score 843.5; DB 2; Length 178;
Best Local Similarity 97.8%; Pred. No. 6.8e-32;
Matches 174; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

```
Qy   11 NTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGKKIHQNNG  70
           |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db    1 NTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGKKIHQNNG  60

Qy   71 LDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKREKIDAAKKCSETFTNKLKEKHTD-L 129
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 LDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKREKIDAAKKCSETFTNKLKEKHTDSF 120

Qy  130 GKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKELTSPVV 187
           ||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db  121 GKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKELTSPVV 178
```

RESULT 12

ID   Q9R7B4   PRELIMINARY;   PRT;   175 AA.
AC   Q9R7B4;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C (FRAGMENT).
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TETS;
RX   MEDLINE=97478003; PubMed=9336916;
RA   Ras N.M., Postic D., Foretz M., Baranton G.;
RT   "Borrelia burgdorferi sensu stricto, a bacterial species 'made in the
RL   U.S.A.'?;
RL   Int. J. Syst. Bacteriol. 47:1112-1117(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TETS;
RA   Marti-Ras N., Postic D., Foretz M., Baranton G.;
RL   Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U91792; AAB81889.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
FT   NON_TER   1   1
FT   NON_TER   175   175
SQ   SEQUENCE   175 AA; 18573 MW; EC059E7B03AC3250 CRC64;

Query Match 44.6%; Score 829; DB 2; Length 175;
Best Local Similarity 97.6%; Pred. No. 3.1e-31;
Matches 166; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

```
Qy    1 MACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKA  60
         ::||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db    6 ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA  65

Qy   61 IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKREKIDAAKKCSETFTN 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   66 IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKREKIDAAKKCSETFTN 125

Qy  121 KLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKA 170
         |||||||||||||||||||||||| ||||||||||||||||||||||||
Db  126 KLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKA 175
```

RESULT 13

ID   Q9R7A9   PRELIMINARY;   PRT;   163 AA.
AC   Q9R7A9;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C (FRAGMENT).
OS   Borrelia burgdorferi sensu stricto.
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TETS;
RX   MEDLINE=97478003; PubMed=9336916;
RA   Ras N.M., Postic D., Foretz M., Baranton G.;
RT   "Borrelia burgdorferi sensu stricto, a bacterial species 'made in
RL   U.S.A.'?;
RL   Int. J. Syst. Bacteriol. 47:1112-1117(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TETS;
RA   Marti-Ras N., Postic D., Foretz M., Baranton G.;
RL   Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U91801; AAB81898.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
FT   NON_TER   1   1
FT   NON_TER   163   163
SQ   SEQUENCE   163 AA; 17210 MW; 7550141651BD01FF CRC64;

```
Query Match
Best Local Similarity   43.1%;   Score 802;  DB 2;  Length 163;
Matches  160;  Conservative  2;  Mismatches  4;  Pred. No. 4.8e-30;  Indels  0;  Gaps  0;

Qy    2 ACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKAI 61
         :|||||||||||||||||||||||||||:|||||||||||||||||||||||||||||||
Db    1 SCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAI 60

Qy   62 GKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNK 121
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 GKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNK 120

Qy  122 LKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESV 164
         |||||||||||||||||||||||||||:|||:|||||||||||
Db  121 LKEKHTDLGKEGVTDADAKEAILKTNGIKTKGAEELGKLFESV 163
```

```
RESULT 14
ID   Q9AGB1                      PRELIMINARY;   PRT;   165 AA.
AC   Q9AGB1;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OSPC-MODIFIED.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Lawson C.L., Swaminathan S.;
RA   Kumaran D., Eswaramoorthy S., Luft B.J., Koide S., Dunn J.J.,
RT   "Crystal structure of outer surface protein C (OspC) from the Lyme
RT   disease spirochete, Borrelia burgdorferi.";
RL   EMBO J. 0:0-0(2001).
DR   EMBL; AF337548; AAK21289.1; -.
FT   VARIANT      61     61       M -> I.
SQ   SEQUENCE   165 AA;   17833 MW;   56E01536D22F61BF CRC64;
```

```
Query Match
Best Local Similarity   42.3%;   Score 786;  DB 2;  Length 165;
Matches  161;  Conservative  2;  Mismatches  1;  Indels  0;  Gaps  0;
Pred. No. 2.6e-29;

Qy  215 VKGPNLTEISKKITESNAVLLAVKEVETLLTSIDELAKAIGKKIKNDVSLDNEADHNGSL 274
         :|||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||
Db    1 MKGPNLTEISKKITESNAVLLAVKEVETLLTSIDELAKAIGKKIRNDVSLDNEADHNGSL 60

Qy  275 ISGAYLISNLITKKISAIKDSGELKAEIEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAK 334
         :|||||| ||:|||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 MSGAYLISTLITKKISAIKDSGELKAEIEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAK 120

Qy  335 KAILKTNNDKTKGADELEKLFESVKNLSRAAKEMLTNSVKELTS 378
         ||||||||||||||||||||||||||||||||||||||||||||
Db  121 KAILKTNNDKTKGADELEKLFESVKNLSRAAKEMLTNSVKELTS 164
```

```
RESULT 15
ID   Q44720                      PRELIMINARY;   PRT;   211 AA.
AC   Q44720;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OSPC.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=25015;
RX   MEDLINE=95154673; PubMed=7851744;
RA   Stevenson B., Barthold S.W.;
RT   "Expression and sequence of outer surface protein C among North
RT   American isolates of Borrelia burgdorferi.";
RL   FEMS Microbiol. Lett. 124:367-372(1994).
DR   EMBL; U04282; AAC45540.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
SQ   SEQUENCE   211 AA;   22603 MW;   63984BA6D8743ED5 CRC64;
```

```
Query Match
Best Local Similarity   41.7%;   Score 774.5;  DB 2;  Length 211;
Matches  158;  Conservative  14;  Mismatches  22;  Indels  1;  Gaps  1;
Pred. No. 1.1e-28;

Qy    1 MACNNSGKDGN-TSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAK 59
         : |||||||| |  |||||||||||||||:||||:|||||||||||||:||||:|:|| |
Db   17 ISCNNSGKDGNAASTNPADESVKGPNLTEISKKIITDSNAVLLAVKEVGALITSIDELATK 76

Qy   60 AIGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAVKKCSEFT 119
         |||||||||||||||||||||||||||||:|||| ||:|||| ::|||:||||:||:|:|
Db   77 AIGKKIHQNNGLDTENNHNGSLLAGAYAISTLITQKLGGLKNEELKEKIAAVKKCSEEFT 136

Qy  120 NKLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKRAAKEMLANSV 179
         |||| ||:|||:  |::||||:|||::  |||:|||||:|||: ||||||||||||||:||
Db  137 NKLKSSHTELGKQDAQDDDAKKAILRTHNTKDKGAEELDKLFKAVENLSKAAKEMLSNSV 196

Qy  180 KELTSPVVAESPKKP 194
         |||||||||||||||
Db  197 KELTSPVVAESPKKP 211
```

Search completed: March 18, 2002, 10:10:53
Job time: 976 sec

THIS PAGE BLANK (USPTO)

OM protein - protein search, using sw model

```
                    GenCore version 4.5
                    Copyright (c) 1993 - 2000 Compugen Ltd.

Run on:             March 18, 2002, 09:54:30 ; Search time 118.14 Seconds
                                                (without alignments)
                                                236.377 Million cell updates/sec

Title:              1850
Perfect score:      1850
Sequence:           1 MACNNSGKDGNTSANSADES.........AVENLAKAAKEMLANSVKEL 377

Scoring table:      BLOSUM62
                    Gapop 10.0 , Gapext 0.5
Searched:           522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters:     522463

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing:    Minimum Match 0%
                    Maximum Match 100%
                    Listing first 45 summaries

Database :          A_Geneseq_1101.*
```

Database	
1:	/SIDS8/gcgdata/geneseq/geneseqp/AA1980.DAT:*
2:	/SIDS8/gcgdata/geneseq/geneseqp/AA1981.DAT:*
3:	/SIDS8/gcgdata/geneseq/geneseqp/AA1982.DAT:*
4:	/SIDS8/gcgdata/geneseq/geneseqp/AA1983.DAT:*
5:	/SIDS8/gcgdata/geneseq/geneseqp/AA1984.DAT:*
6:	/SIDS8/gcgdata/geneseq/geneseqp/AA1985.DAT:*
7:	/SIDS8/gcgdata/geneseq/geneseqp/AA1986.DAT:*
8:	/SIDS8/gcgdata/geneseq/geneseqp/AA1987.DAT:*
9:	/SIDS8/gcgdata/geneseq/geneseqp/AA1988.DAT:*
10:	/SIDS8/gcgdata/geneseq/geneseqp/AA1989.DAT:*
11:	/SIDS8/gcgdata/geneseq/geneseqp/AA1990.DAT:*
12:	/SIDS8/gcgdata/geneseq/geneseqp/AA1991.DAT:*
13:	/SIDS8/gcgdata/geneseq/geneseqp/AA1992.DAT:*
14:	/SIDS8/gcgdata/geneseq/geneseqp/AA1993.DAT:*
15:	/SIDS8/gcgdata/geneseq/geneseqp/AA1994.DAT:*
16:	/SIDS8/gcgdata/geneseq/geneseqp/AA1995.DAT:*
17:	/SIDS8/gcgdata/geneseq/geneseqp/AA1996.DAT:*
18:	/SIDS8/gcgdata/geneseq/geneseqp/AA1997.DAT:*
19:	/SIDS8/gcgdata/geneseq/geneseqp/AA1998.DAT:*
20:	/SIDS8/gcgdata/geneseq/geneseqp/AA1999.DAT:*
21:	/SIDS8/gcgdata/geneseq/geneseqp/AA2000.DAT:*
22:	/SIDS8/gcgdata/geneseq/geneseqp/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	% Query Match	Length	DB	ID	Description
1	1850	100.0	377	22	AAB62713	Borrelia sp chimer
2	1825	98.6	400	22	AAB62739	Borrelia sp chimer
3	1796.5	97.1	373	22	AAB62711	Borrelia sp chimer
4	1793.5	96.9	397	22	AAB62729	Borrelia sp chimer
5	1614.5	87.3	378	22	AAB62712	Borrelia sp chimer
6	1589.5	85.9	401	22	AAB62738	Borrelia sp chimer
7	1561	84.4	374	22	AAB62710	Borrelia sp chimer
8	1558	84.2	398	22	AAB62728	Borrelia sp chimer
9	1519.5	82.1	369	22	AAB62716	Borrelia sp chimer
10	1516.5	82.0	393	22	AAB62731	Borrelia sp chimer
11	1480.5	80.0	367	22	AAB62719	Borrelia sp chimer
12	1477.5	79.9	391	22	AAB62735	Borrelia sp chimer
13	1467	79.3	384	22	AAB62726	Borrelia sp chimer
14	1465	79.2	560	22	AAB62724	Borrelia sp chimer
15	1464	79.1	408	22	AAB62737	Borrelia sp chimer
16	1284	69.4	370	22	AAB62715	Borrelia sp chimer
17	1281	69.2	394	22	AAB62730	Borrelia sp chimer
18	1259.5	68.1	369	22	AAB62714	Borrelia sp chimer
19	1256.5	67.9	369	22	AAB62732	Borrelia sp chimer
20	1245	67.3	368	22	AAB62718	Borrelia sp chimer
21	1242	67.1	392	22	AAB62734	Borrelia sp chimer
22	1225.5	66.2	368	22	AAB62717	Borrelia sp chimer
23	1217.5	65.8	391	22	AAB62736	Borrelia sp chimer
24	1183.5	64.0	378	22	AAB62725	Borrelia sp chimer
25	1175.5	63.5	401	22	AAB62733	Borrelia sp chimer
26	1164	62.9	410	22	AAB62740	Borrelia sp chimer
27	1163	62.9	386	22	AAB62727	Borrelia sp chimer
28	972.5	52.6	466	16	AAR57540	B31 outer surface
29	954	51.6	209	22	AAB62720	Borrelia sp chimer
30	945	51.1	210	16	AAB62733	Outer surface prot
31	943	51.0	210	15	AAW11935	Borrelia IP2 OspC
32	939	50.8	192	22	AAR60886	B31 outer surface
33	929	50.2	192	22	AAB62721	Borrelia OspC
34	928	50.2	210	16	AAR75746	B31 Osp-A/antigen
35	915	49.5	466	16	AAR75727	B. burgdorferi str
36	908	49.1	192	22	AAR75739	B31 outer surface
37	876	47.4	193	22	AAB62703	B burgdorferi OspC
38	864	46.7	176	22	AAB62709	B burgdorferi OspC
39	707	38.2	192	15	AAR62772	Borrelia B31 antig
40	694.5	37.5	177	15	AAB62773	Borrelia 297 antig
41	694.5	37.5	177	15	AAR60889	Borrelia 25015 ant
42	692.5	37.4	211	18	AAR60887	Borrelia 25015 Osp
43	692.5	37.4	211	20	AAW41822	B. burgdorferi sen
44	690.5	37.3	193	15	AAW93473	B. burgdorferi Osp
45	685	37.0	212	16	AAR60888	Borrelia ZS7 antig
			193	15	AAW11934	Outer surface prot

ALIGNMENTS

```
RESULT   1
AAB62713
ID   AAB62713 standard; Protein; 377 AA.
XX
AC   AAB62713;
XX
DT   03-APR-2001  (first entry)
XX
DE   Borrelia sp chimeric ospc protein SEQ ID NO: 30.
XX
KW   Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
XX
PN   WO200078966-A1.
XX
PD   28-DEC-2000.
XX
PF   19-JUN-2000; 2000WO-US16915.
XX
PR   18-JUN-1999;  99US-0140042.
XX
PA   (UYNY ) UNIV NEW YORK STATE RES FOUND.
PA   (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI   Dattwyler RJ,  Seinost G,  Dykhuizen D,  Luft BJ,  Gomes-Solecki M;
DR   WPI; 2001-050113/06.
DR   N-PSDB; AAF29017.
XX
PT   Compositions of OspC polypeptides from strains of Borrelia which cause
PT   Lyme disease are used to immunize animals and detect immune responses
PT   to Lyme disease -
```

RESULT 2
AAB62739

AC   AAB62739;

DT   03-APR-2001 (first entry)

DE   Borrelia sp chimeric ospC protein SEQ ID NO: 82.

KW   Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.

OS   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.

PN   WO200078966-A1.

PD   28-DEC-2000.

PR   19-JUN-2000; 2000WO-US16915.

PR   18-JUN-1999; 99US-0140042.

PA   (UYNY) UNIV NEW YORK STATE RES FOUND.
PA   (BROO-) BROOK BIOTECHNOLOGIES INC.

PI   Dattwyler RJ, Seinost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;

XX
PS   Claim 43; Page 86-87; 160pp; English.
XX
XX
CC   The present invention provides compositions comprising ospC proteins and
CC   chimeric ospC proteins from members of the Borrelia genus. These may be
CC   Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC   vaccines against Borrelia infection, which is spread by ticks and leads
CC   to Lyme disease.
XX
SQ   Sequence   377 AA;

Query Match   100.0%; Score 1850; DB 22; Length 377;
Best Local Similarity   100.0%; Pred. No. 1.6e-124;
Matches   377; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 MACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKA 60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 macnnsgkdgntsansadesvkgpnlteinkkitdsnavllavkeveallssideiaaka 60

Qy   61 IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKOKLDGLKNEGLKEKIDAAKKCSETFTN 120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 igkkihqnngldtennhngsllagayaistlikokldglknegklkekidaakkcsetftn 120

Qy 121 KLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 klkekhtdlgkegvtdadakeailkangtktkgaeelgklfesvevlskaakemlansvk 180

Qy 181 ELTSPVVAESPKKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEI 240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 eltspvvaespkkpsmvnnsgkdgntsansadesvkgpnlteiskkitesnavvlavkei 240

Qy 241 ETLLASIDELATKAIGKKIQQNGGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLK 300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 241 etllasidelatkaigkkiqqngglaveaghngtllagaytisklitqkldglknseklk 300

Qy 301 EKIENAKKCSEDFTKKLEGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVE 360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 301 ekienakkcsedftkklegehaqlgienvtdenakkailitdaakdkgaaelekifkave 360

Qy 361 NLAKAAAKEMLANSVKEL 377
         ||||||||||||||||||
Db 361 nlakaaakemlansvkel 377

RESULT 3
AAB62711

ID   AAB62711 standard; Protein; 373 AA.
XX
AC   AAB62711;
XX
DT   03-APR-2001 (first entry)
XX
DE   Borrelia sp chimeric ospC protein SEQ ID NO: 26.
XX
KW   Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
XX
PN   WO200078966-A1.
XX
PD   28-DEC-2000.
XX
PR   19-JUN-2000; 2000WO-US16915.
PR   18-JUN-1999; 99US-0140042.

DR   WPI; 2001-050113/06.
DR   N-PSDB; AAF29043.
XX
PT   Compositions of OspC polypeptides from strains of Borrelia which cause
PT   Lyme disease are used to immunize animals and detect immune responses
PT   to Lyme disease -
XX
PS   Claim 43; Page 153; 160pp; English.
XX
XX
CC   The present invention provides compositions comprising ospC proteins and
CC   chimeric ospC proteins from members of the Borrelia genus. These may be
CC   Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC   vaccines against Borrelia infection, which is spread by ticks and leads
CC   to Lyme disease.
XX
SQ   Sequence   400 AA;

Query Match   98.6%; Score 1825; DB 22; Length 400;
Best Local Similarity   98.9%; Pred. No. 1.1e-122;
Matches   372; Conservative   2; Mismatches   2; Indels   0; Gaps   0;

Qy   2 ACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKAI 61
         :|||||||||||||||||||||||||:|||||||:|||||||||||||||||||||||||
Db   25 scnnsgkdgntsansadesvkgpnlteiskkitdsnavllavkeveallssideiaakai 84

Qy   62 GKKIHQNNGLDTENNHNGSLLAGAYAISTLIKOKLDGLKNEGLKEKIDAAKKCSETFTNK 121
         |||||||||:|||||||||||||||||||||||||||||||||||||||||||||:|||
Db   85 gkkihqnngldteynhngsllagayaistlikqkldglknegklkekidaakkcsetftnk 144

Qy 122 LKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 181
         |||||||||||||||||||||||:|||||||||||||||||||||||||||||||||||
Db 145 lkekhtdlgkegvtdadakeailktngtktkgaeelgklfesvevlskaakemlansvke 204

Qy 182 LTSPVVAESPKKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIE 241
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 205 ltspvvaespkkpsmvnnsgkdgntsansadesvkgpnlteiskkitesnavvlavkeie 264

Qy 242 TLLASIDELATKAIGKKIQONGGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKE 301
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 265 tllasidelatkaigkkiqongggaveaghngtllagaytisklitqkldglknseklke 324

Qy 302 KIENAKKCSEDFTKKLEGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVEN 361
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 325 kienakkcsedftkklegehaqlgienvtdenakkailitdaakdkgaaelekifkaven 384

Qy 362 LAKAAAKEMLANSVKEL 377
         |||||||||||||||||
Db 385 lakaaakemlansvkel 400

RESULT 4

AAB62729

ID   AAB62729 standard; Protein; 397 AA.

XX

AC   AAB62729;

XX

DT   03-APR-2001 (first entry)

XX

DE   Borrelia sp chimeric ospc protein SEQ ID NO: 62.

XX

KW   Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.

OS   Chimeric - Borrelia sp.

OS   Chimeric - Borrelia sp.

XX

PN   WO200078966-A1.

XX

PD   28-DEC-2000.

XX

PF   19-JUN-2000; 2000WO-US16915.

XX

PR   18-JUN-1999; 99US-0140042.

XX

PA   (UYNY) UNIV NEW YORK STATE RES FOUND.

PA   (BROO-) BROOK BIOTECHNOLOGIES INC.

XX

PI   Dattwyler RJ, Seinost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;

XX

DR   WPI; 2001-050113/06.

DR   N-PSDB; AAF29033.

XX

PT   Compositions of OspC polypeptides from strains of Borrelia which cause

PT   Lyme disease are used to immunize animals and detect immune responses

PT   to Lyme disease -

XX

PS   Claim 43; Page 125-126; 160pp; English.

XX

CC   The present invention provides compositions comprising ospc proteins and

CC   chimeric ospc proteins from members of the Borrelia genus. These may be

CC   Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as

CC   vaccines against Borrelia infection, which is spread by ticks and leads

CC   to Lyme disease.

XX

SQ   Sequence   397 AA;

Query Match   96.9%; Score 1793.5; DB 22; Length 397;
Best Local Similarity   97.9%; Pred. No. 1.9e-120;
Matches 368; Conservative   3; Mismatches   2; Indels   3; Gaps   1;

Qy   2 ACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKAI 61
               :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2 scnnsgkdgntsansadesvkgpnlteinkkitdsnavllavkeveallssideiaakai 61

Qy   62 GKKIHQNNGLDTENNHNGSLLAGAYAIASTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNK 121
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   62 gkkihqnngldtennhngsllagayaiastlikqkldglkneglkekidaakkcsetftnk 121

Qy   122 LKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 181
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   122 lkekhtdlgkegvtdadakeailkangtktkgaeelgklfesvevlskaakemlansvke 181

Qy   182 LTSPVVAESPKKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIE 241
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   182 ltspvvaespkkpsmvnnsgkdgntsansadesvkgpnlteiskkitesnavvlavkeie 241

Qy   242 TLLASIDELATKAIGKKIQQNGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKE 301
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   242 tllasidelatkaigkkiqqngglaveaghngtllagaytisklitqkldglknseklke 301

Qy   302 KIENAKKCSEDFTKKLEGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVEN 361
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   302 kienakkcsedftkklegehaqlgienvtdenakkailitdaakdkgaaeleklfkaven 361

Qy   362 LAKAAKEMLANSVKEL 377
               ||||||||||||||||
Db   362 lakaakemlansvkel 377

Db   382 lakaakemlansvkel 397

RESULT 5

AAB62712

ID   AAB62712 standard; Protein; 378 AA.

XX

AC   AAB62712;

XX

DT   03-APR-2001 (first entry)

XX

DE   Borrelia sp chimeric ospc protein SEQ ID NO: 28.

XX

Query Match   97.1%; Score 1796.5; DB 22; Length 373;
Best Local Similarity   98.1%; Pred. No. 1e-120;
Matches 369; Conservative   2; Mismatches   2; Indels   3; Gaps   1;

Qy   2 ACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKAI 61
               :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 acnnsgkdgntsansadesvkgpnlteinkkitdsnavllavkeveallssideiaakai 60

Qy   62 GKKIHQNNGLDTENNHNGSLLAGAYAIASTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNK 121
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 gkkihqnngldteynhngsllagayaistlikqkldglkneglkekidaakkcsetftnk 120

Qy   122 LKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 181
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 lkekhtdlgkegvtdadakeailktngtktkgaeelgklfesvevlskaakemlansvke 180

Qy   182 LTSPVVAESPKKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIE 241
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 ltspvvaes---pamvnnsgkdgntsansadesvkgpnlteiskkitesnavvlavkeie 237

Qy   242 TLLASIDELATKAIGKKIQQNGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKE 301
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   238 tllasidelatkaigkkiqqngglaveaghngtllagaytisklitqkldglknseklke 297

Qy   302 KIENAKKCSEDFTKKLEGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVEN 361
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   298 kienakkcsedftkklegehaqlgienvtdenakkailitdaakdkgaaeleklfkaven 357

Qy   362 LAKAAKEMLANSVKEL 377
               ||||||||||||||||
Db   358 lakaakemlansvkel 373

Left column

```
KW   Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
XX
XX
PN   WO200078966-A1.
XX
PD   28-DEC-2000.
XX
PF   19-JUN-2000; 2000WO-US16915.
XX
PR   18-JUN-1999; 99US-0140042.
XX
PN
PA   (UYNY ) UNIV NEW YORK STATE RES FOUND.
PA   (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI   Dattwyler RJ, Seinost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
PT   WPI; 2001-050113/06.
PT   N-PSDB; AAF29016.
XX

CC   Compositions of OspC polypeptides from strains of Borrelia which cause
CC   Lyme disease are used to immunize animals and detect immune responses
CC   to Lyme disease -

The present invention provides compositions comprising ospc proteins and
chimeric ospc proteins from members of the Borrelia genus. These may be
Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
vaccines against Borrelia infection, which is spread by ticks and leads
to Lyme disease.

Claim 43; Page 83-84; 160pp; English.

SQ   Sequence   378 AA;

Query Match            87.3%;
Best Local Similarity  87.8%;
Matches  331;  Conservative  20;  Mismatches  25;  Indels  1;  Gaps  1;

Score 1614.5;  Pred. No. 1.1e-107;  DB 22;  Length 378;

Qy    1 MACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKVEVEALLSSIDEIAAKA  60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 macnnsgkdgntsansadesvkgpnlteinkkitdsnavllavkveveallssideiaaka  60

Qy   61 IGKKIHQNNGLDTENNINGSLLAGAYAISTLLIKQKLDGLKNEGLKEKIDAAKKCSETFTN 120
         ||||||||||| ||||||:|||||||||||||||||||||||||||||||||||||||||
Db   61 igkkihqnngldtennhngsllagayaistllikqkldglkneglkekidaakkcsetftn 120

Qy  121 KLKEKHTDLGKEGVTDADAKERAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 180
         |||||||||||||||||||||:|||||||||||||||||||||||||||||||||||||
Db  121 klkekhtdlgkegvtdadakeailkangtktkgaeelgklfesvevlskaakemlansvk  180

Qy  181 ELTSPVVAESPKKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEI  240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 eltspvvaespkkpsmvnnsgkdgntsansadesvkgpnlteiskkitesnavvlavkei  240

Qy  241 ETLLASIDELATKAIGKKIQONGGIAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLK  300
         |||| ||||| |:||||::|: : |::|||||| |:|| :||| |::|:||||::|:||
Db  241 etlltsidela-kaigkkindvslidneadhngslisgaylisnliitkkisalkdsgelk 299

Qy  301 EKIENAKKCSEDFTKLEGEHAQLGIENYTDENAKKAILITDAAKDKGAAELEKLFRAVE   360
         :::|:||||:|:||:| |||:|:| |||| ::|||| |:|:::|:||:|||||| |:|
Db  300 aeiekakkcseeftaklkgehtdlgkegvtddnakkaliktnndktkgadelekflfesvk 359

Qy  361 NLAKAAKEMLANSVKEL 377
         ||:|||||||:|||||
Db  360 nlskaakemltnsvkel 376
```

Right column

```
AC   AAB62738;
XX
DT   03-APR-2001 (first entry)
XX
DE   Borrelia sp chimeric ospc protein SEQ ID NO: 80.
XX
KW   Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
XX
PN   WO200078966-A1.
XX
PD   28-DEC-2000.
XX
PF   19-JUN-2000; 2000WO-US16915.
XX
PR   18-JUN-1999; 99US-0140042.
XX
PN
PA   (UYNY ) UNIV NEW YORK STATE RES FOUND.
PA   (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI   Dattwyler RJ, Seinost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
PT   WPI; 2001-050113/06.
PT   N-PSDB; AAF29042.
XX

CC   Claim 43; Page 150-151; 160pp; English.

The present invention provides compositions comprising ospc proteins and
chimeric ospc proteins from members of the Borrelia genus. These may be
Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
vaccines against Borrelia infection, which is spread by ticks and leads
to Lyme disease.

Compositions of OspC polypeptides from strains of Borrelia which cause
Lyme disease are used to immunize animals and detect immune responses
to Lyme disease -

SQ   Sequence   401 AA;

Query Match            85.9%;
Best Local Similarity  86.7%;
Matches  326;  Conservative  22;  Mismatches  27;  Indels  1;  Gaps  1;

Score 1589.5;  Pred. No. 7.2e-106;  DB 22;  Length 401;

Qy    2 ACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKVEVEALLSSIDEIAAKAI  61
         :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:
Db   25 scnnsgkdgntsansadesvkgpnlteinkkitdsnavllavkveveallssideiaaka   84

Qy   62 GKKIHQNNGLDTENNINGSLLAGAYAISTLLIKQKLDGLKNEGLKEKIDAAKKCSETFTNK 121
         ||||||||||| ||:|:||||||||||||||||||||||||||||||||||||||||||:
Db   85 gkkihqnngldteynhngsllagayaistllikqkldglkneglkekidaakkcsetftnk 144

Qy  122 LKEKHTDLGKEGVTDADAKERAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 181
         |||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||
Db  145 lkekhtdlgkegvtdadakeailkangtktkgaeelgklfesvevlskaakemlansvke  204

Qy  182 LTSPVVAESPKKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIE  241
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|
Db  205 ltspvvaespkkpsmvnnsgkdgntsansadesvkgpnlteiskkitesnavvlavkeve  264

Qy  242 TLLLASIDELATKAIGKKIQONGGIAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKE 301
         ||| ||||| |:||||::|: : |::|||||| |:|| :||| |::|:||||::|:||:
Db  265 tlltsidela-kaigkkindvslidneadhngslisgaylisnliitkkisalkdsgelka 323

Qy  302 KIENAKKCSEDFTKLEGEHAQLGIENYTDENAKKAILITDAAKDKGAAELEKLFRAVEN  361
         :::|:||||:|:||:| |||:|:| |||| ::|||| |:|:::|:||:|||||| |:||
Db  324 eiekakkcseeftaklkgehtdlgkegvtddnakkaliktnndktkgadelekflfesvkn 383

Qy  362 LAKAAKEMLANSVKEL 377
         |:|||||||:|||||
Db  384 lskaakemltnsvkel 399
```

Bottom (left)

```
RESULT 6
AAB62738
ID   AAB62738 standard; Protein; 401 AA.
XX
```

RESULT 7

AAB62710

ID   AAB62710 standard; Protein; 374 AA.
XX
AC   AAB62710;
XX
DT   03-APR-2001 (first entry)
XX
DE   Borrelia sp chimeric ospc protein SEQ ID NO: 24.
XX
KW   Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
XX
XX
PN   WO200078966-A1.
XX
PD   28-DEC-2000.
XX
PF   19-JUN-2000; 2000WO-US140042.
XX
PR   18-JUN-1999; 99US-0140042.
XX
PT   Compositions of OspC polypeptides from strains of Borrelia which cause
PT   Lyme disease are used to immunize animals and detect immune responses
XX
PA   (UYNY) UNIV NEW YORK STATE RES FOUND.
PA   (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI   Dattwyler RJ, Seinost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
DR   WPI; 2001-050113/06.
DR   N-PSDB; AAF29014.
XX
PS   Claim 43; Page 78-79; 160pp; English.
XX
CC   The present invention provides compositions comprising ospc proteins and
CC   chimeric ospc proteins from members of the Borrelia genus. These may be
CC   Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC   vaccines against Borrelia infection, which is spread by ticks and leads
CC   to Lyme disease.
XX
SQ   Sequence   374 AA;

Query Match   84.4%; Score 1561; DB 22; Length 374;
Best Local Similarity   85.9%; Pred. No. 7.1e-104;
Matches 323; Conservative   22; Mismatches 27; Indels   4; Gaps   2;

Qy   2 ACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKAI   61
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 acnnsgkdgntsansadesvkgpnlteinkkitdsnavllavkeveallssideiaakai   60

Qy   62 GKKIHQNNGLDTENNHNGSLLAGAYAISTLIKOKLDGLKNEGLKEKIDAAKKCSETFTNK 121
            ||||||||||||| |||||||||||||||||:|||||||:||||||||||||||||||||
Db   61 gkkihqnngldteynhngsllagayaistlikqkidglkneglkekidaakkcsetftnk 120

Qy   122 LKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 181
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 lkekhtdlgkegvtdadakeailktngtktkgaeelgklfesvevlskaakemlansvke 180

Qy   182 LTSPVVAESPKKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIE 241
            |||||||| :|||||||||||||||||||||||||||||||| |||||||||||||:|:|
Db   181 ltspvvaes--pamvnnsgkdgntsansadesvkgpnlteiskkitesnavvlavkeve 237

Qy   242 TLLASIDELATKAIGKKITQQNGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKE 301
            ||||||||||| ||||||||| |||||||||||||||:|| ::||||||||||:|:|:
Db   238 tiltsidela-kaigkkikndvsidneedhngsllsgaylisnlitkkisalkdsgelka 296

Qy   302 KIENAKKCSEDFTKKLEGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVEN 361
            ||| |||||||||| |||| :::|| || |||||||:||||||||||||||||||| |||
Db   297 eiekakkcseeftaklkgehtdlgkegvtddnakkailktndkctkgadelekflesvkn 356

Qy   362 LAKAAKEMLANSVKEL 377
            |:||||||||||||:|
Db   357 lskaakemltnsvkel 372

RESULT 8

AAB62728

ID   AAB62728 standard; Protein; 398 AA.
XX
AC   AAB62728;
XX
DT   03-APR-2001 (first entry)
XX
DE   Borrelia sp chimeric ospc protein SEQ ID NO: 60.
XX
KW   Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
XX
XX
PN   WO200078966-A1.
XX
PD   28-DEC-2000.
XX
PF   19-JUN-2000; 2000WO-US140042.
XX
PR   18-JUN-1999; 99US-0140042.
XX
PT   Compositions of OspC polypeptides from strains of Borrelia which cause
PT   Lyme disease are used to immunize animals and detect immune responses
XX
PA   (UYNY) UNIV NEW YORK STATE RES FOUND.
PA   (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI   Dattwyler RJ, Seinost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
DR   WPI; 2001-050113/06.
DR   N-PSDB; AAF29032.
XX
PS   Claim 43; Page 123; 160pp; English.
XX
CC   The present invention provides compositions comprising ospc proteins and
CC   chimeric ospc proteins from members of the Borrelia genus. These may be
CC   Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC   vaccines against Borrelia infection, which is spread by ticks and leads
CC   to Lyme disease.
XX
SQ   Sequence   398 AA;

Query Match   84.2%; Score 1558; DB 22; Length 398;
Best Local Similarity   85.6%; Pred. No. 1.3e-103;
Matches 322; Conservative   23; Mismatches 27; Indels   4; Gaps   2;

Qy   2 ACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKAI   61
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   25 scnnsgkdgntsansadesvkgpnlteinkkitdsnavllavkeveallssideiaakai   84

Qy   62 GKKIHQNNGLDTENNHNGSLLAGAYAISTLIKOKLDGLKNEGLKEKIDAAKKCSETFTNK 121
            ||||||||||||| |||||||||||||||:|||||||:||||||||||||||||||||
Db   85 gkkihqnngldteynhngsllagayaistlikqkidglkneglkekidaakkcsetftnk 144

Qy   122 LKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 181
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   145 lkekhtdlgkegvtdadakeailktngtktkgaeelgklfesvevlskaakemlansvke 204

Qy   182 LTSPVVAESPKKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIE 241
            |||||||| :|||||||||||||||||||||||||||||||| ||||||||||||||:|:
Db   205 ltspvvaes--pamvnnsgkdgntsansadesvkgpnlteiskkitesnavvlavkeve 261

RESULT 9

ID AAB62716 standard; Protein; 369 AA.
XX
AC AAB62716;
XX
DT 03-APR-2001 (first entry)
XX
DE Borrelia sp chimeric ospC protein SEQ ID NO: 36.
XX
KW Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS Chimeric - Borrelia sp.
OS Chimeric - Borrelia sp.
XX
PN WO200078966-A1.
XX
PD 28-DEC-2000.
XX
PF 19-JUN-2000; 2000WO-US016915.
XX
PR 18-JUN-1999; 99US-0140042.
XX
PA (UYNY) UNIV NEW YORK STATE RES FOUND.
PA (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI Dattwyler RJ, Seinost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
DR WPI; 2001-050113/06.
DR N-PSDB; AAF29020.
XX
PT Compositions of OspC polypeptides from strains of Borrelia which cause
PT Lyme disease are used to immunize animals and detect immune responses
PT to Lyme disease.
XX
PS Claim 43; Page 94-95; 160pp; English.
XX
CC The present invention provides compositions comprising ospC proteins and
CC chimeric ospC proteins from members of the Borrelia genus. These may be
CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC vaccines against Borrelia infection, which is spread by ticks and leads
CC to Lyme disease.
XX
SQ Sequence 369 AA;

Query Match 82.1%; Score 1519.5; DB 22; Length 369;
Best Local Similarity 86.2%; Pred. No. 6.5e-101;
Matches 326; Conservative 10; Mismatches 31; Indels 11; Gaps 4;

Qy 2 ACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKAI 61
|||| ||: ||||||||||||||||||||| |||||||||||||||||: ||||||||||
Db 1 acnnsgkdgntsansadesvkgpnlteiskkitdsnavllavkeveallssidel-akai 59

Qy 62 GKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLIDGLK-NEGLKEKIDAAKKCSETFTN 120
||||| : ||||||||||||||||||||||||: |||||| :|||||||||||| |:
Db 60 gkkihqngldnenernesllagaytistlitqklsklngseglkekiaaakkcseefst 119

Qy 121 KLKEKHTDLGLGKEGVTDADAKEAILKANGT-KTKGAEELGKLFESVEVLSKAAKEMLANSV 179
||||: |||||||||||||||||||:||||| |||: ||||||||||||||||||||||||
Db 121 klkekhtdlgkegvtdadakeailkangt-ktkgaeelgklfesvevlskaakemlansv 179

RESULT 10

ID AAB62731 standard; Protein; 393 AA.
XX
AC AAB62731;
XX
DT 03-APR-2001 (first entry)
XX
DE Borrelia sp chimeric ospC protein SEQ ID NO: 66.
XX
KW Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS Chimeric - Borrelia sp.
OS Chimeric - Borrelia sp.
XX
PN WO200078966-A1.
XX
PD 28-DEC-2000.
XX
PF 19-JUN-2000; 2000WO-US016915.
XX
PR 18-JUN-1999; 99US-0140042.
XX
PA (UYNY) UNIV NEW YORK STATE RES FOUND.
PA (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI Dattwyler RJ, Seinost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
DR WPI; 2001-050113/06.
DR N-PSDB; AAF29035.
XX
PT Compositions of OspC polypeptides from strains of Borrelia which cause
PT Lyme disease are used to immunize animals and detect immune responses
PT to Lyme disease.
XX
PS Claim 43; Page 131; 160pp; English.
XX
CC The present invention provides compositions comprising ospC proteins and
CC chimeric ospC proteins from members of the Borrelia genus. These may be
CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC vaccines against Borrelia infection, which is spread by ticks and leads
CC to Lyme disease.
XX
SQ Sequence 393 AA;

Query Match 82.0%; Score 1516.5; DB 22; Length 393;
Best Local Similarity 86.0%; Pred. No. 1.2e-100;
Matches 325; Conservative 11; Mismatches 31; Indels 11; Gaps 4;

Qy 2 ACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKAI 61
|||| ||||||||||||||||||||||||||||||||||||||||||||: ||||||||
Db 25 scnnsgkdgntsansadesvkgpnlteiskkitdsnavllavkeveallssidel-akai 83

Qy 60 GKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLIDGLK-NEGLKEKIDAAKKCSETFTN 120
||| |: |||||||||||||||||||||||||: |||||| :|||||||||||| ||
Db 62 GKKIHQNGLDTENNHNGSLLAGAYAISTLIKQKLIDGLIK-NEGLKEKIDAAKKCSETFTN 120

Qy 180 keltspvvhg------nnsgkdgntsansadesvkgpnlteiskkitesnavvlavke 239

Qy 240 IETLLIASIDELATKAIGKKIQONGGLAVEAGHNGTLLAGAYTISKLLTQKLIDGLK 299
|||||::|||||||||||||||||:|:||||||||||||||||||||||||||:|||||
Db 232 ietllasidelatkaigkkiiqqngglaveaghngtllagaytisklitqklidglknsek 291

Qy 300 KEKIENAKKCSEDFTKKLEGEHAQLGIENVTDENAKKAILITDAAKDKGAAAELEKLFKAV 359
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 292 kekienakkcsedftkklegehaqlgienvtdenakkailitdaakdkgaaelekifkav 351

Qy 360 ENLAKAAKEMLANSVKEL 377
||||||||||||||||||
Db 352 enlakaakemlansvkel 369

```
Qy   62  GKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLK-NEGLKEKIDAAKKCSETFTN  120
         |||| |   |  | |:|:||||||||||||||||| :||  ||:||||||| ||||:||| |:
Db   84  gkkikndgsldneanrnesllagaytistitqkiskiqnsegikekiaaakkcseefst   143

Qy  121  KLKEKHTDLGKEGVTDADAKEAILKANGT-KTKGAEELGKLFESVEVLSKAAKEMLANSV  179
         ||| :|  | |::||||:||:||||||:|  |.|| |||:||||||||||||||||||||
Db  144  klkdnhaqlgigqvtdenakkallkandagkdkqveeleklsgslesleslskaakemlansv 203

Qy  180  KELTSPVVAESPKKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKE  239
         |||||||| |       ||||||||||||||||||||||||||||||||||||||||||
Db  204  keltspvvhg------nnsgkdgntsansadesvkgpnlteiskkitesnavvlavke    255

Qy  240  IETLLASIDELATKAIGKKIQONGGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKL  299
         ||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||
Db  256  ietllasidelatkaigkkiqqngglaveaghngtllagaytiskliitqkldglknsekl 315

Qy  300  KEKIENAKKCSEDFTKKLEGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAV  359
         |||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||
Db  316  kekienakkcsedftkklegehaqlgienvtdenakkaliitdaakdkgaaeleklfkav   375

Qy  360  ENLAKAAAKEMLANSVKEL  377
         ||||| |||||||||||||
Db  376  enlakaakemlansvkel   393
```

RESULT 11
AAB62719

ID   AAB62719 standard; Protein; 367 AA.
XX
AC   AAB62719;
XX
DT   03-APR-2001 (first entry)
XX
DE   Borrelia sp chimeric ospc protein SEQ ID NO: 42.
XX
KW   Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
XX
PN   WO20007896966-A1.
XX
PD   28-DEC-2000.
XX
PF   19-JUN-2000; 2000WO-US16915.
XX
PR   18-JUN-1999; 99US-0140042.
XX
PA   (UYNY) UNIV NEW YORK STATE RES FOUND.
PA   (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI   Dattwyler RJ, Seinost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
DR   WPI; 2001-050113/06.
DR   N-PSDB; AAF29023.
XX
PT   Claim 43; Page 102-103; 160pp; English.
PT   Compositions of OspC polypeptides from strains of Borrelia which cause
PT   Lyme disease are used to immunize animals and detect immune responses
PT   to Lyme disease -
XX
CC   The present invention provides compositions comprising ospc proteins and
CC   chimeric ospc proteins from members of the Borrelia genus. These may be
CC   Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC   vaccines against Borrelia infection, which is spread by ticks and leads
CC   to Lyme disease.
XX
SQ   Sequence 367 AA;

Query Match 80.0%; Score 1480.5; DB 22; Length 367;

Best Local Similarity 82.4%; Pred. No. 3.9e-98;
Matches 310; Conservative 23; Mismatches 34; Indels 9; Gaps 2;

```
Qy    2  ACNNSCKDGNTSANSADESVAGPNLTEINKIITDSNAVLLAVKEVEALLSSIDEIAAKAI   61
         ||||| |||| :|||||||| :|||||| ::|| |:.|:||| ||:: ||| ||::|:|||
Db    1  acnnsgkdgnasansadesvkgpnlteiskkitesnavvlavkevetllasidelatkai   60

Qy   62  GKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNK  121
         |||| |||| |   |||| || ||||| :||::| |:|||:| ||||||:|: ||||||||
Db   61  gkki-gnngleangsknts11sgayaisdliaeklnvlkneelkekidtakqscsteftnk 119

Qy  122  LKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKEIE 181
         ||: |:|| ::|:||::  :|||| .| |.||||||||||||||||||||::|: |::||
Db  120  lksehavlgldnitdddnaqrailkkhankdkgaaeleklfkavenlskaaqdtiknavke  179

Qy  182  LTSPVVAESPKKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIE  241
         |||| |         ||||||||||||||||||||||||||||||||||||||||||||
Db  180  ltspivhg------nnsgkdgntsansadesvkgpnlteiskkitesnavvlavkele    231

Qy  242  TLLASIDELATKAIGKKIQONGGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKE  301
         |||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||
Db  232  tllasidelatkaigkkiqqngglaveaghngtllagaytiskliitqkldglknseklke 291

Qy  302  KIENAKKCSEDFTKKLEGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVEN  361
         |||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||
Db  292  kienakkcsedftkklegehaqlgienvtdenakkaliitdaakdkgaaeleklfkaven   351

Qy  362  LAKAAAKEMLANSVKEL  377
         ||||| |||||||||||
Db  352  lakaakemlansvkel   367
```

RESULT 12
AAB62735

ID   AAB62735 standard; Protein; 391 AA.
XX
AC   AAB62735;
XX
DT   03-APR-2001 (first entry)
XX
DE   Borrelia sp chimeric ospc protein SEQ ID NO: 74.
XX
KW   Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
XX
PN   WO20007896966-A1.
XX
PD   28-DEC-2000.
XX
PF   19-JUN-2000; 2000WO-US16915.
XX
PR   18-JUN-1999; 99US-0140042.
XX
PA   (UYNY) UNIV NEW YORK STATE RES FOUND.
PA   (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI   Dattwyler RJ, Seinost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
DR   WPI; 2001-050113/06.
DR   N-PSDB; AAF29039.
XX
PT   Claim 43; Page 142; 160pp; English.
PT   Compositions of OspC polypeptides from strains of Borrelia which cause
PT   Lyme disease are used to immunize animals and detect immune responses
PT   to Lyme disease -
XX
CC   The present invention provides compositions comprising ospc proteins and
CC   chimeric ospc proteins from members of the Borrelia genus. These may be
CC   Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC   vaccines against Borrelia infection, which is spread by ticks and leads

Left column (RESULT 13)

```
CC    to Lyme disease.
XX
SQ    Sequence   391 AA;
XX

Query Match          79.98;   Score 1477.5;  DB 22;  Length 391;
Best Local Similarity  82.2%;  Pred. No. 7e-98;
Matches 309;  Conservative  24;  Mismatches  34;  Indels  9;  Gaps  2;

Qy     2 ACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKAI 61
          ::|||||||| :||||||||||||||| :|| |||||:||| || :|| ||||||| |:|
Db    25 scnnsgkdgnasansadesvkgpnlteiskkitesnavvlavkeveetllasideiatkai 84

Qy    62 GKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLIKNEGLKEKIDAAKKCSEFTFNK 121
          ||| ||:||  :|||| ::||||||||||| :|:|||||:| |||||||||||||||||
Db    85 gkki-qnngleangsknltslisgayaisdliaeklnvlkneelkekidtakqcsteftnk 143

Qy   122 LKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 181
          ||||| ||:|||||||||||| :
Db   144 lksehavlgidnltddnaqrailkhandkgaaelekifkavenlskaaqdtilknavke 203

Qy   182 LTSPVVAESPKKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESSNAVVLAVKEIE 241
          |||:|  :                   :||||||||||||||||||||||||||||||||
Db   204 ltspivhg------nnsgkdgntsansadesvkgpnlteiskkitesnavvlavkeie 255

Qy   242 TLIASIDELATKAIGKKIQONGGLAVEAGHNGTLLAGAYTISKLITOKLDGLKNSEKLKE 301
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   256 tl1asidelatkaigkkiqqngglaveaghngtllagaytiskiitqkldglknseklke 315

Qy   302 KIENAKKCSEDFTKKLEGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVEN 361
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   316 kienakkcsedftkklegehaqlgienvtdenakkailitdaakdkgaaelekifkaven 375

Qy   362 LAKAAKEMLANSVKEL 377
          |||||||||||||||
Db   376 lakaakemlansvkel 391
```

```
RESULT 13
AAB62726
ID    AAB62726 standard; Protein; 384 AA.
XX
AC    AAB62726;
XX
DT    03-APR-2001  (first entry)
XX
KW    Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS    Chimeric - Borrelia sp.
OS    Chimeric - Borrelia sp.
XX
DE    Borrelia sp chimeric ospc protein SEQ ID NO: 56.
XX
PN    WO200078966-A1.
XX
PD    28-DEC-2000.
XX
PF    19-JUN-2000; 2000WO-US16915.
XX
PR    18-JUN-1999;  99US-0140042.
XX
PA    (UYNY ) UNIV NEW YORK STATE RES FOUND.
PA    (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI    Dattwyler RJ, Seinost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
DR    WPI; 2001-050113/06.
DR    N-PSDB; AAP29030.
XX
PT    Compositions of OspC polypeptides from strains of Borrelia which cause
PT    Lyme disease are used to immunize animals and detect immune responses
PT    to Lyme disease -
```

Right column (RESULT 14)

```
PS    Claim 43; Page 117-118; 160pp; English.
XX
CC    The present invention provides compositions comprising ospc proteins and
CC    chimeric ospC proteins from members of the Borrelia genus. These may be
CC    Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC    vaccines against Borrelia infection, which is spread by ticks and leads
CC    to Lyme disease.
XX
SQ    Sequence   384 AA;
XX

Query Match          79.3%;   Score 1467;  DB 22;  Length 384;
Best Local Similarity  81.4%;  Pred. No. 3.9e-97;
Matches 305;  Conservative  30;  Mismatches  36;  Indels  4;  Gaps  3;

Qy     2 ACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKAI 61
          |::||||||||||||||||||||||||| :|||||||||||||||||||||||||||||
Db     1 acnnsgkdgntsansadesvkgpnlteiskkitdsnavllavkeveeallssideiaakai 60

Qy    62 GKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLIKNEGLKEKIDAAKKCSEFTFNK 121
          ||||  :||||| :||||| |||||||||||||:|:|||||:| |||||||||||||||||
Db    61 gkkihqngngldteynhngsllagayaistlikqkidglikneglkekidaakkcsetftnk 121

Qy   122 LKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 181
          |||||||||||| :||||||||| ||||||||||||||||||||||||||||||||||||
Db   121 lkekhtdlgkegvedadakeailktngtktkgaeelgklfesvevlskaakemlansvke 181

Qy   182 LTSPVVAESPKKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESSNAVVLAVKEIE 241
          ||||||||||||:| :||||||||||||||||||||||||||||||||| ::|||||||||
Db   181 ltspvvaespkkpkpfhgnnsg--gdsastnpdesakgpnltviskaitdsnafllavkeve 238

Qy   242 TLIASIDELATKAIGKKIQONGGLAVEAGHNGTLLAGAYTISKLITOKLDGLKNSEKLKE 301
          ||||||||| :| ::||| :: |   :||||||||||||||||||||:| ||:|||||||
Db   239 allssidel-skaigkkikndgtldneanrnesllagayeiskltqklsvl-nseelkk 296

Qy   302 KIENAKKCSEDFTKKLEGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVEN 361
          |||||||||||||:|||||||:|| :||| |  :|| :|||||||||||||||||::|:
Db   297 kikeakdcsqkfttklkdshaelgiqsvqddnakkailkthgtkdkgakeleelfksles 356

Qy   362 LAKAAKEMLANSVKEL 377
          |:||||| |:|||||
Db   357 lskaaqqaltnsvkel 372
```

```
RESULT 14
AAB62724
ID    AAB62724 standard; Protein; 560 AA.
XX
AC    AAB62724;
XX
DT    03-APR-2001  (first entry)
XX
KW    Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS    Chimeric - Borrelia sp.
OS    Chimeric - Borrelia sp.
XX
DE    Borrelia sp chimeric ospc protein SEQ ID NO: 52.
XX
PN    WO200078966-A1.
XX
PD    28-DEC-2000.
XX
PF    19-JUN-2000; 2000WO-US16915.
XX
PR    18-JUN-1999;  99US-0140042.
XX
PA    (UYNY ) UNIV NEW YORK STATE RES FOUND.
PA    (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI    Dattwyler RJ, Seinost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
DR    WPI; 2001-050113/06.
```

Left column

```
DR   N-PSDB; AAF29028.
XX
PT   Compositions of OspC polypeptides from strains of Borrelia which cause
PT   Lyme disease are used to immunize animals and detect immune responses
PT   to Lyme disease -
XX
PS   Claim 43; Page 112-113; 160pp; English.
XX
CC   The present invention provides compositions comprising ospc proteins and
CC   chimeric ospC proteins from members of the Borrelia genus. These may be
CC   Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC   vaccines against Borrelia infection, which is spread by ticks and leads
CC   to Lyme disease.
XX
SQ   Sequence   560 AA;

Query Match             79.2%;  Score 1465;  DB 22;  Length 560;
Best Local Similarity   81.0%;
Matches 306;  Conservative  23;  Mismatches  45;  Indels  4;  Gaps  2;

Qy      1  MACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKA   60
           |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db      1  macnnsgkdgntsansadesvkgpnlteiskkitdsnavllavkeveallssideiaaka   60

Qy     61  IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGCLKNEGLKEKIDAARKCSETFTN  120
           ||||||||||||| ||||||||||||||| ||||||||||||| ||||||||||||||||
Db     61  igkkihqnngldteynhngsllagayaistllkqkldglknegikekidaakkcsetftn  120

Qy    121  KLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEEELGKLFESVEVLSKAAKEMLANSVK  180
           ||||||||||||||||||||||| ||||||||||| |||||||||||||||||||||||||
Db    121  klkekhtdlgkegvtdadakeailktngtktkgaeelgklfesvevlskaakemlansvk  180

Qy    181  ELTSPVVAESPKKPSMVNNSGKDGNT-SANSADESVKGPNLTEISKKITESNAVVLAVKE   239
           |||||||| | :|  :||| ||||||  |||||||||||||||||||||||| |||||||
Db    181  eltspvvaes---pamgsnsgkggdsastnpadesakgpnlteiskkitdsnafvlavke  237

Qy    240  IETLLASIDELATKAIGKKIQONGGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKL   299
           :|:|: ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    238  vetivlsidelakkalgqkidnnglaalnqngsllagayaistliteklsklknleel   297

Qy    300  KEKIENAKKCSEDFTKKLEGEHAQLGIENVTDENAKNKAILTDAAKDKGAAELEKLFKAV  359
           |:||||||||:|::|||||:||:::  |  |||  |:|||:: :|||||||||:  |:|
Db    298  kteiakakkcseeftnklksghadigkqdacddhakeaailkthattcdkgakefkdlfesv  357

Qy    360  ENLAKAAKEMLANSVKEL  377
           |||| | |||||||||||
Db    358  eglikaaqvalnsvkel  375
```

Right column

```
PA   (UYNY ) UNIV NEW YORK STATE RES FOUND.
PA   (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI   Dattwyler RJ, Seinost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
```

```
DR   WPI; 2001-050113/06.
DR   N-PSDB; AAF29041.
XX
PT   Compositions of OspC polypeptides from strains of Borrelia which cause
PT   Lyme disease are used to immunize animals and detect immune responses
PT   to Lyme disease -
XX
PS   Claim 43; Page 147-148; 160pp; English.
XX
CC   The present invention provides compositions comprising ospc proteins and
CC   chimeric ospC proteins from members of the Borrelia genus. These may be
CC   Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC   vaccines against Borrelia infection, which is spread by ticks and leads
CC   to Lyme disease.
XX
SQ   Sequence   408 AA;

Query Match             79.1%;  Score 1464;  DB 22;  Length 408;
Best Local Similarity   81.1%;
Matches 305;  Conservative  31;  Mismatches  36;  Indels  4;  Gaps  3;

Qy      2  ACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKAI   61
           :|||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db     25  scnnsgkdgntsansadesvkgpnlteiskkitdsnavllavkeveallssideiaakai   84

Qy     62  GKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAARKCSETFTNK  121
           ||||||||||| |||||||||||||||| |||||||||||| |||||||||||||||||
Db     85  gkkihqnngldteynhngsllagayaistliqkqkldglknegikekidaakkcsetftnk  144

Qy    122  LKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE  181
           |||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db    145  lkekhtdlgkegvtdadakeailktngtktkgaeelgklfesvevlskaakemlansvke  204

Qy    182  LTSPVVAESPKKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIE  241
           |||||||||||:|| ||||  :||||||||||||||||||||||||||||||||||||||
Db    205  ltspvvaespkkpfhgnnsg--ggdsastnpadesakgpnlteiskkitdsnafilavkeve  262

Qy    242  TLLASIDELATKAIGKKIQONGGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLK   301
           |||||||||| :|| |||| || | ||| |||||||||||||||||||||||||||||||
Db    263  allssidel-skaigkkindgtldneanrneslilagayeiskiltqklsvl-nseelk   320

Qy    302  KIENAKKCSEDFTKKLEGEHAQLGIENVTDENAKNKAILTDAAKDKGAAELEKLFKAVEN  361
           |: |||||||||:|| || ||:  |  | | ::||  |::|:: :|||||||:| |:
Db    321  kikeakdcsqkftcklkdshaelgiqsvgddnakkaailkthgtkdkgakeleelfksies  380

Qy    362  LAKAAKEMLANSVKEL  377
           |||| | |||||||||
Db    381  lskaaqaltnsvkel  396
```

Bottom-left (database entry)

```
                    15
   AAB62737
ID        AAB62737   standard;  protein;  408 AA.
XX
AC        AAB62737;
XX
DT        03-APR-2001  (first entry)
XX
DE        Borrelia sp chimeric ospC protein SEQ ID NO: 78.
XX
KW        ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS        Borrelia sp.
OS        Chimeric - Borrelia sp.
XX
PN        WO200078966-A1.
XX
PD        28-DEC-2000.
XX
PF        19-JUN-2000; 2000WO-US16915.
XX
PR        18-JUN-1999; 99US-0140042.
XX
```

Search completed: March 18, 2002, 09:54:31
Job time: 329 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 09:55:32 ; Search time 55.5 Seconds
                                 (without alignments)
                                 152.860 Million cell updates/sec

Title: US-09-596-746A-30
Perfect score: 1850
Sequence: 1 MACNNSGKDGNTSANSADES..............AVENLAKAAKEMLANSVKEL 377

Scoring table: BLOSUM62
Gap open 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   Issued_Patents_AA:*
          1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*
          2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*
          3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:*
          4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:*
          5: /cgn2_6/ptodata/2/1aa/PCTUS_COMB.pep:*
          6: /cgn2_6/ptodata/2/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	Match DB	DB ID	Description
1	972.5	52.6	466	4	US-08-235-836C-110	Sequence 110, App
2	945	51.1	210	4	US-08-158-353-3	Sequence 3, Appli
3	938	50.7	210	4	US-08-209-603E-15	Sequence 15, Appli
4	933	50.4	209	4	US-09-196-293-15	Sequence 15, Appli
5	932	50.4	210	4	US-08-235-836C-30	Sequence 30, Appli
6	929	50.2	588	4	US-08-235-836C-122	Sequence 122, Appl
7	915	49.5	466	4	US-08-235-836C-107	Sequence 107, App
8	685	37.0	212	1	US-08-158-353-2	Sequence 2, Appli
9	670	36.2	212	1	US-08-031-295-2	Sequence 2, Appli
10	670	36.2	212	1	US-07-903-580-2	Sequence 2, Appli
11	631	34.1	212	4	US-08-158-353-4	Sequence 4, Appli
12	629	34.0	212	1	US-09-196-293-11	Sequence 11, Appl
13	629	34.0	212	1	US-08-209-603E-11	Sequence 11, Appli
14	626	33.8	212	4	US-08-235-836C-34	Sequence 34, Appl
15	614.5	33.2	209	1	US-08-235-836C-32	Sequence 32, Appl
16	592.5	32.0	207	4	US-08-235-836C-36	Sequence 36, Appl
17	171	9.2	194	4	US-09-364-083-2	Sequence 2, Appli
18	168	9.1	630	4	US-08-973-462-9	Sequence 9, Appli
19	164.5	8.9	1164	4	US-08-923-992A-10	Sequence 10, Appl
20	162.5	8.8	1164	4	US-08-923-992A-2	Sequence 2, Appli
21	160	8.6	1786	4	US-08-973-462-8	Sequence 8, Appli
22	158	8.5	1098	4	US-08-923-992A-8	Sequence 8, Appli
23	154.5	8.4	1128	4	US-08-923-992A-6	Sequence 6, Appli
24	151.5	8.2	1104	4	US-08-923-992A-4	Sequence 4, Appli
25	148	8.0	3111	2	US-08-460-309-4	Sequence 4, Appli
26	148	8.0	3111	2	US-08-125-077-4	Sequence 4, Appli
27	147.5	8.0	2482	1	US-08-328-254-6	Sequence 6, Appli
28	146.5	7.9	688	3	US-09-141-047-8	Sequence 8, Appli
29	146.5	7.9	3248	3	US-08-353-700-1	Sequence 1, Appli
30	146.5	7.9	3248	5	PCT-US95-16216-1	Sequence 1, Appli
31	145	7.8	1388	2	US-08-685-576-1	Sequence 1, Appli
32	143.5	7.8	1388	2	US-08-685-576-4	Sequence 4, Appli
33	143	7.7	414	5	PCT-US93-03077-3	Sequence 3, Appli
34	143	7.7	1093	5	PCT-US93-03077-1	Sequence 1, Appli
35	140.5	7.6	1196	2	US-08-144-121-4	Sequence 4, Appli
36	140.5	7.6	1196	2	US-08-735-893-4	Sequence 4, Appli
37	140	7.6	1196	2	US-08-470-260-5	Sequence 5, Appli
38	140	7.6	1147	3	US-08-471-491-5	Sequence 5, Appli
39	140	7.6	3289	2	US-08-466-662-5	Sequence 5, Appli
40	140	7.6	1147	4	US-08-477-451-2	Sequence 2, Appli
41	139.5	7.5	396	1	US-08-430-024-2	Sequence 2, Appli
42	139.5	7.5	396	1	US-08-782-009-2	Sequence 2, Appli
43	139.5	7.5	396	3	US-09-017-302-2	Sequence 2, Appli
44	133.5	7.2	1354	3	US-08-685-871-2	Sequence 2, Appli
45	132	7.1	2285	4	US-09-308-375-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-08-235-836C-110
Sequence 110, Application US/08235836C
Patent No. 6248562

GENERAL INFORMATION:
; APPLICANT: Dunn, John J.
; APPLICANT: Luft, Benjamin J.
; TITLE OF INVENTION: No. 6248562e1 Chimeric Proteins Comprising
; TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor
; NUMBER OF SEQUENCES: 144
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brookhaven National Laboratory
; STREET:
; CITY: Upton
; STATE: NY
; COUNTRY: USA
; ZIP: 11973
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/235,836C
; FILING DATE: 29-APR-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/148,191
; FILING DATE: 01-11-93
; ATTORNEY/AGENT INFORMATION:
; NAME: Bogosian, Margaret C.
; REGISTRATION NUMBER: 25,324
; REFERENCE/DOCKET NUMBER: BNL93-28A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 282-7338
; TELEFAX: (516) 282-3729
; INFORMATION FOR SEQ ID NO:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 466 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-235-836C-110

Qy   1 MACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKA 60

          1 MACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKA 60

Query Match   52.6%;   Score 972.5;   DB 4;   Length 466;
Best Local Similarity   62.9%;   Pred. No. 5.8e-66;
Matches   224;   Conservative   23;   Mismatches   74;   Indels   35;   Gaps   8;

```
Db    17  ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA  76
          ::|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    61  IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN  120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    77  IGKKIHQNNGLDTEYNHNGSLLAGRYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN  136
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   121  KLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK  180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   137  KLKAKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK  196
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   181  ELTSPVVAESPKKPSMVNNGSKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEI  240
          |||||||||||||||||
Db   197  ELTSPVVAESPKKPKQ-NVSSLD---EKNSVSVDLPGEMKVLVSKEKNKDGKY------- 245
          ::|:.:|:                  :||:
Qy   241  ETTLASIDELATKAIGKKIQQNGGL-AVERGHNGTLLA---------------GAYTIS  283
          |:|:.:|:                      :|:              :|
Db   246 -DLIATVDKLELKGTSDKNNGSGVLEGVKADKSKVKLTISDDLGQTTLEVFKEDGKTLVS 304
          :|
Db   284  KLITQKLDGLKNSEKLKEKIENAKKCSEDFTKKLEGEHAQ-LGIENVTDENAKKAI  338
          ||:.:
Db   305  KKVTSK-DKSSTEEKFNEKGE----VSEKIITRADGTRLEYTGIKSDGSGKAKEVL  355
```

```
RESULT 2
US-08-158-353-3
Sequence 3, Application US/08158353
Patent No. 5620862

GENERAL INFORMATION:
APPLICANT: Padula, Steven J.
TITLE OF INVENTION: Methods for Diagnosing Early Lyme
                    Disease
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
  ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
  STREET: Two Militia Drive
  CITY: Lexington
  STATE: MA
  COUNTRY: USA
  ZIP: 02173
COMPUTER READABLE FORM:
  MEDIUM TYPE: Floppy disk
  COMPUTER: IBM PC compatible
  OPERATING SYSTEM: PC-DOS/MS-DOS
  SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/158,353
  FILING DATE:
  CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
  NAME: Carroll, Alice O.
  REGISTRATION NUMBER: 33,542
  REFERENCE/DOCKET NUMBER: UCT93-05
TELECOMMUNICATION INFORMATION:
  TELEPHONE: 617-861-6240
  TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
  LENGTH: 210 amino acids
  TYPE: amino acid
  STRANDEDNESS: single
  TOPOLOGY: linear
  MOLECULE TYPE: protein
```

```
Qy   1  MACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKA  60
        :.||||||||||||||||||||||||||||.|||||.|||||||||||||||||:|||||
Db  17  ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA  76
```

```
Query Match            51.1%; Score 945; DB 1; Length 210;
Best Local Similarity  97.9%; Pred. No. 2.3e-64;
Matches 190;  Conservative 3; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy    61  IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN  120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    77  IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN  136
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   121  KLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK  180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   137  KLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK  196
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   181  ELTSPVVAESPKKP  194
          ||||||||||||||
Db   197  ELTSPVVAESPKKP  210
```

```
RESULT 3
US-08-209-603E-15
Sequence 15, Application US/08209603E
Patent No. 6248538

GENERAL INFORMATION:
APPLICANT: FUCHS, RENATE
APPLICANT: WILSKE, BETTINA
APPLICANT: PREAC-MURSIC, VERA
APPLICANT: MOTZ, MANFRED
APPLICANT: SOUTSCHECK, ERWIN
TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PROTEINS
TITLE OF INVENTION: FROM BORRELIA BURGDORFERI
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
  ADDRESSEE: BROOKS HAIDT HAFFNER & DELAHUNTY
  STREET: 99 PARK AVENUE
  CITY: NEW YORK
  STATE: NY
  COUNTRY: USA
  ZIP: 10016
COMPUTER READABLE FORM:
  MEDIUM TYPE: 3.5" FLOPPY DISC
  COMPUTER: AT&T - IBM COMPATIBLE
  OPERATING SYSTEM: MS-DOS Version 6.2
  SOFTWARE: ASCII
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/209,603E
  FILING DATE: 10-MAR-1994
  CLASSIFICATION: 436
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: PCT/EP90/02282
  FILING DATE: 21-DEC-1990
  APPLICATION NUMBER: US 07/862,535
  FILING DATE: 19-JUN-1992
ATTORNEY/AGENT INFORMATION:
  NAME: ROBINSON, WILLIAM R.
  REGISTRATION NUMBER: 27,224
  REFERENCE/DOCKET NUMBER: LKR-9217-A
TELECOMMUNICATION INFORMATION:
  TELEPHONE: (212) 697-3355
  TELEFAX: (212) 557-5635
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
  LENGTH: 210
  TYPE: AMINO ACID
  TOPOLOGY: LINEAR
  MOLECULE TYPE: PROTEIN
  DESCRIPTION:
  HYPOTHETICAL: N/A
  ANTI-SENSE: N/A
  FRAGMENT TYPE: N/A
  ORIGINAL SOURCE:
  ORGANISM: B. BURGDORFERI
  IMMEDIATE SOURCE:
  LIBRARY: DSM 5662
  POSITION IN GENOME: N/A
  FEATURE:
  IDENTIFICATION METHOD: amino acid analysis
  PUBLICATION INFORMATION: N/A
```

US-08-209-603E-15

Query Match   50.7%;   Score 938;   DB 4;   Length 210;
Best Local Similarity   97.4%;   Pred. No. 7.9e-64;
Matches 189;   Conservative   3;   Mismatches   2;   Indels   0;   Gaps   0;

```
Qy   1  MACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKA  60
        ::||||||||||||||||||||||||||||::||||||||||||:|||||||||||||||
Db  17  ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVLEVEALLSSIDEIAAKA  76

Qy  61  IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN  120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  77  IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN  136

Qy 121  KLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK  180
        ||||||||||||||||||||||||||:|||||||||||||||||||||||||||||||||
Db 137  KLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK  196

Qy 181  ELTSPVVAESPKK  194
        |||||||||||||
Db 197  ELTSPVVAESPKKP  210
```

RESULT 4
US-09-196-293-15
Sequence 15, Application US/0919293
Patent No. 6183755
GENERAL INFORMATION:
;   APPLICANT: Fuchs, Renate
;   APPLICANT: Motz, Manfred
;   APPLICANT: Soutscheck, Erwin
;   APPLICANT: Wilske, Bettina
;   APPLICANT: Preac-Mursic, Vera
;   TITLE OF INVENTION: Active proteins from Borrelia
;   TITLE OF INVENTION: burgdorferi
;   FILE REFERENCE: 738. 001US2
;   CURRENT APPLICATION NUMBER: US/09/196,293
;   CURRENT FILING DATE: 1998-11-19
;   EARLIER APPLICATION NUMBER: US 08/209,603
;   EARLIER FILING DATE: 1994-03-10
;   EARLIER APPLICATION NUMBER: US 07/862,535
;   EARLIER FILING DATE: 1992-06-19
;   EARLIER APPLICATION NUMBER: WO PCT/EP90/02282
;   EARLIER FILING DATE: 1990-12-21
;   EARLIER APPLICATION NUMBER: DE P39 42 728.5
;   EARLIER FILING DATE: 1989-12-22
;   EARLIER APPLICATION NUMBER: DE P40 18 988.0
;   EARLIER FILING DATE: 1990-06-13
;   NUMBER OF SEQ ID NOS: 16
;   SEQ ID NO 15
;   SOFTWARE: FastSEQ for Windows Version 4.0
;   LENGTH: 209
;   TYPE: PRT
;   ORGANISM: Borrelia burgdorferi
US-09-196-293-15

Query Match   50.4%;   Score 933;   DB 4;   Length 209;
Best Local Similarity   97.9%;   Pred. No. 1.9e-63;
Matches 188;   Conservative   3;   Mismatches   1;   Indels   0;   Gaps   0;

```
Qy   1  MACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKA  60
        ::||||||||||||||||||||||||||||::||||||||||||:|||||||||||||||
Db  17  ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVLEVEALLSSIDEIAAKA  76

Qy  61  IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN  120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  77  IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN  136

Qy 121  KLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK  180
        ||||||||||||||||||||||||||:|||||||||||||||||||||||||||||||||
Db 137  KLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK  196

Qy 181  ELTSPVVAESPKK  194
        |||||||||||||
Db 197  ELTSPVVAESPKKP  210
```

RESULT 5
US-08-235-836C-30
Sequence 30, Application US/08235836C
Patent No. 6248562
GENERAL INFORMATION:
;   APPLICANT: Dunn, John J.
;   APPLICANT: Luft, Benjamin J.
;   TITLE OF INVENTION: No. 6248562e1 Chimeric Proteins Comprising
;   TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor
;   NUMBER OF SEQUENCES: 144
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Brookhaven National Laboratory
;   STREET:
;   CITY: Upton
;   STATE: NY
;   COUNTRY: USA
;   ZIP: 11973
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/235,836C
;   FILING DATE: 29-APR-1994
;   CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 08/148,191
;   FILING DATE: 01-11-93
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Bogosian, Margaret C.
;   REGISTRATION NUMBER: 25,324
;   REFERENCE/DOCKET NUMBER: BNL93-28A
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (516) 282-7338
;   TELEFAX: (516) 282-3729
INFORMATION FOR SEQ ID NO: 30:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 210 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-08-235-836C-30

```
Qy 181  ELTSPVVAESPK  192
        ||||||||||||
Db 197  ELTSPVVAESPK  208
```

Query Match   50.4%;   Score 932;   DB 4;   Length 210;
Best Local Similarity   96.9%;   Pred. No. 2.2e-63;
Matches 188;   Conservative   3;   Mismatches   3;   Indels   0;   Gaps   0;

```
Qy   1  MACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKA  60
        ::||||||||||||||||||||||||||||::||||||||||||:|||||||||||||||
Db  17  ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVLEVEALLSSIDEIAAKA  76

Qy  61  IGKKIHQNNGLDTEYNHNGSLLAGRYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN  120
        |||||||||||||:|||||||||:||||||||||||||||||||||||||||||||||||
Db  77  IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN  136

Qy 121  KLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK  180
        ||||||||||||||||||||||||||:|||||||||||||||||||||||||||||||||
Db 137  KLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK  196

Qy 181  ELTSPVVAESPKK  194
        |||||||||||||
Db 197  ELTSPVVAESPKKP  210
```

RESULT 6

US-08-235-836C-122

; Sequence 122, Application US/08235836C
; Patent No. 6248562
; GENERAL INFORMATION:
; APPLICANT: Dunn, John J.
; APPLICANT: Luft, Benjamin J
; TITLE OF INVENTION: No. 62485262e1 Chimeric Proteins Comprising
; TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor
; NUMBER OF SEQUENCES: 144
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brookhaven National Laboratory
; STREET:
; CITY: Upton
; STATE: NY
; COUNTRY: USA
; ZIP: 11973
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/235,836C
; FILING DATE: 29-APR-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/148,191
; FILING DATE: 01-11-93
; ATTORNEY/AGENT INFORMATION:
; NAME: Bogosian, Margaret C.
; REGISTRATION NUMBER: 25,324
; REFERENCE/DOCKET NUMBER: BNL93-28A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 282-7338
; TELEFAX: (516) 282-3729
; INFORMATION FOR SEQ ID NO: 122:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 588 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-235-836C-122

Query Match   50.2%;   Score 929; DB 4; Length 588;
Best Local Similarity   89.7%; Pred. No. 1.5e-62;
Matches 191; Conservative 6; Mismatches 16; Indels   0; Gaps 0;

Qy   4 NNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGK 63
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   286 NNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGK 345

Qy   64 KIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNKLK 123
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   346 KIHQNNGLDTEYNHNGSLLAGRYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNKLK 405

Qy   124 EKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKELT 183
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   406 EKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKELT 465

Qy   184 SPVVAESPKPSPSMVNNSGKDGNTSANSADESVK 216
           ||||||||||| : |||||:||:
Db   466 SPVVAESPKKPGTMAQYNQMHMLSNKSASQNVR 498

RESULT 7
US-08-235-836C-107
; Sequence 107, Application US/08235836C
; Patent No. 6248562
; GENERAL INFORMATION:
; APPLICANT: Dunn, John J.
; APPLICANT: Luft, Benjamin J.
; TITLE OF INVENTION: No. 62485262e1 Chimeric Proteins Comprising

---

; TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor
; NUMBER OF SEQUENCES: 144
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brookhaven National Laboratory
; STREET:
; CITY: Upton
; STATE: NY
; COUNTRY: USA
; ZIP: 11973
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/235,836C
; FILING DATE: 29-APR-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/148,191
; FILING DATE: 01-11-93
; ATTORNEY/AGENT INFORMATION:
; NAME: Bogosian, Margaret C.
; REGISTRATION NUMBER: 25,324
; REFERENCE/DOCKET NUMBER: BNL93-28A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 282-7338
; TELEFAX: (516) 282-3729
; INFORMATION FOR SEQ ID NO: 107:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 466 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-235-836C-107

Query Match   49.5%;   Score 915; DB 4; Length 466;
Best Local Similarity   97.4%; Pred. No. 1.3e-61;
Matches 186; Conservative 1; Mismatches 4; Indels   0; Gaps 0;

Qy   4 NNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGK 63
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   276 NNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGK 335

Qy   64 KIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNKLK 123
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   336 KIHQNNGLDTEYNHNGSLLAGRYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNKLK 395

Qy   124 EKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKELT 183
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   396 AKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKELT 455

Qy   184 SPVVAESPKKP 194
           |||||||||||
Db   456 SPVVAESPKKP 466

RESULT 8
US-08-158-353-2
; Sequence 2, Application US/08158353
; Patent No. 5620862
; GENERAL INFORMATION:
; APPLICANT: Padula, Steven J.
; TITLE OF INVENTION: Methods for Diagnosing Early Lyme
; TITLE OF INVENTION: Disease
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: MA
; COUNTRY: USA

ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/158,353
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Carroll, Alice O.
REGISTRATION NUMBER: 33,542
REFERENCE/DOCKET NUMBER: UCT93-05
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-861-6240
TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 212 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-158-353-2

Query Match   37.0%; Score 685; DB 1; Length 212;
Best Local Similarity   74.1%; Pred. No. 1e-44;
Matches 146; Conservative 19; Mismatches 28; Indels 4; Gaps 3;

```
Qy   1 MACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKA 60
         ::|||||||:|||||||||||||||||||||||:||||||||||||||||:|:||| ||
Db  17 ISCNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDEVAKKA 76

Qy  61 IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKN-EGLKEKIDAAKKCSETFT 119
         ||:::|||||||||||||:|||||||||||::|||||||||: :|||||||:||||||||
Db  77 IGNLIAQ-NGLNAGANQNGSLLAGAYVISTLIAEKLDGLKNSEELKEKIEDAKKCNKAFT 135

Qy 120 NKLKEKHTDLG--KEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLAN 177
         :||| | ||||  :|||||||||||||||||:|||:|:|||||||:|||:|:||||
Db 136 DKLKSSHAELGIANGAATDANRKAAILKTNGTKDKGAQELEKLFESVKNLSKAAQETLNN 195

Qy 178 SVKELTSPVVAESPKKP 194
         |||||||||||::|||
Db 196 SVKELTSPVVAENPKKP 212
```

RESULT 9
US-08-031-295-2
Sequence 2, Application US/08031295
Patent No. 5530103
GENERAL INFORMATION:
APPLICANT: LIVEY, Ian
APPLICANT: DORNER, Freidrich
TITLE OF INVENTION: METHOD AND COMPOSITION FOR THE
TITLE OF INVENTION: PREVENTION OF LYME DISEASE
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/031,295
FILING DATE: 19930312
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/903,580
FILING DATE: 25-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/824,161
FILING DATE: 22-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/727,245
FILING DATE: 11-JUL-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/142 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 212 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-031-295-2

Query Match   36.2%; Score 670; DB 1; Length 212;
Best Local Similarity   73.0%; Pred. No. 1.4e-43;
Matches 143; Conservative 15; Mismatches 36; Indels 2; Gaps 2;

```
Qy   1 MACNNSGKDGNT-SANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAK 59
         ::||||| |||| |||||||||||||||||||||||||||:|||||||||:||||||:|
Db  17 ISCNNSGKGGDSASTNPADESARGPNLTEISKKITDSNAFVLAVKEVETLVSSIDELATK 76

Qy  60 AIGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEG-LKEKIDAAKKCSETF 118
         |||||:||||| :|||||||||||||:|||||| |: ||:| | ||:|||:||||||:|
Db  77 AIGKKIQONNGLGANADKNGSLLAGAYAISTLITEKLKALKNSGELKAKIEDAKKCSEDF 136

Qy 119 TNKLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANS 178
         |:|| |:  ||: | |:||:|||||||:|||||||||||| | :: : :|||||:: |:|
Db 137 TKKLAGHAQLGIDGAVDNDSKEAILKTNGTKTKGAEELVKLSESSVASLSKAAQEASANS 196

Qy 179 VKELTSPVVAESPKKP 194
         |||||||||||:|||
Db 197 VKELTSPVVAEIPKKP 212
```

RESULT 10
US-07-903-580-2
Sequence 2, Application US/07903580
Patent No. 6221363
GENERAL INFORMATION:
APPLICANT: LIVEY, Ian
APPLICANT: DORNER, Freidrich
TITLE OF INVENTION: METHOD AND COMPOSITION FOR THE
TITLE OF INVENTION: PREVENTION OF LYME DISEASE
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: Virginia
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/903,580
FILING DATE: 19920625
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/824,161
FILING DATE: 22-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/727,245
FILING DATE: 11-JUL-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/131 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 212 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein

US-07-903-580-2

Query Match   36.2%;   Score 670;   DB 4;   Length 212;
Best Local Similarity 73.0%;   Pred. No. 1.4e-43;
Matches 143; Conservative 15; Mismatches 36; Indels 2; Gaps 2;

Qy   1 MACNNSGKDGNT-SANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAK 59
          ::||||| ||:.|||||||||||||||||:|||||||| :|||||||: |||:|||||
Db   17 ISCNNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETLVSSIDELATK 76

Qy   60 AIGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEG-LKEKIDAAKKCSETF 118
          ||||||:.|||| :|||| ||:||||||||||:.||: || |: |||:|||||||||||
Db   77 AIGKKIQONNGLGANADKNGSLLAGAYAISTLITEKLKALKNSGELKAKIEDAKKCSEDF 136

Qy 119 TNKLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANS 178
          |.|:|: .:|||:|.| |.:|||| :||||||||||||| ||||||||||||||:|| ||
Db 137 TKKLAAGHAQLGIDGATDNDSKEAILKTNGTKTKGAEELVKLSESVASLSKAAQEASANS 196

Qy 179 VKELTSPVVAESPKKP 194
          ||||||||||:|||||
Db 197 VKELTSPVVAETPKKP 212

RESULT 11
US-08-158-353-4
Sequence 4, Application US/08158353
Patent No. 5620862
GENERAL INFORMATION:
APPLICANT: Padula, Steven J.
TITLE OF INVENTION: Methods for Diagnosing Early Lyme
TITLE OF INVENTION: Disease
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Militia Drive
CITY: Lexington
STATE: MA
COUNTRY: USA
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/158,353
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Carroll, Alice O.
REGISTRATION NUMBER: 33,542
REFERENCE/DOCKET NUMBER: UCT93-05

TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-861-6240
TELEFAX: 617-861-9540

INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 212 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-158-353-4

Query Match   34.1%;   Score 631;   DB 1;   Length 212;
Best Local Similarity 69.9%;   Pred. No. 1.3e-40;
Matches 137; Conservative 15; Mismatches 42; Indels 2; Gaps 2;

Qy   1 MACNNSGKDGNT-SANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAK 59
          ::||||| ||:.||||||||||||||||: :||||||: :|||||||: |||:|:|||
Db   17 ISCNNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETLVSSIDELATK 76

Qy   60 AIGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKN-EGLKEKIDAAKKCSETF 118
          ||||||:.|||| :||| ||:||||||||||:.||: || |: :|||:|||||||||:|
Db   77 AIGKIDNNNGLAALNNQNGSLLAGAYAISTLITEKLSKLKNLEELKTEIAKAKKCSEEF 136

Qy 119 TNKLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANS 178
          |.|:|: .:| |||:|.| |.:|||| :||||||||||| ||||||||||||:|| ||
Db 137 TNKLKSGHADLGKODATDDHAKAAILKTHATTDKGAKEFKDLFESVEGLLKAAQVALTNS 196

Qy 179 VKELTSPVVAESPKKP 194
          ||||||||||:|||||
Db 197 VKELTSPVVAETPKKP 212

RESULT 12
US-09-196-293-11
Sequence 11, Application US/09196293
Patent No. 6183755
GENERAL INFORMATION:
APPLICANT: Fuchs, Renate
APPLICANT: Motz, Manfred
APPLICANT: Soutscheck, Erwin
APPLICANT: Wilske, Bettina
APPLICANT: Preac-Mursic, Vera
TITLE OF INVENTION: Active proteins from Borrelia
TITLE OF INVENTION: burgdorferi
FILE REFERENCE: 738.001US2
CURRENT APPLICATION NUMBER: US/09/196,293
CURRENT FILING DATE: 1998-11-19
EARLIER APPLICATION NUMBER: US 09/196,293
EARLIER FILING DATE: 1994-03-10
EARLIER APPLICATION NUMBER: US 08/209,603
EARLIER FILING DATE: 1992-06-19
EARLIER APPLICATION NUMBER: US 07/862,535
EARLIER FILING DATE: 1990-12-21
EARLIER APPLICATION NUMBER: WO PCT/EP90/02282
EARLIER FILING DATE: 1989-12-22
EARLIER APPLICATION NUMBER: DE P39 42 728.5
EARLIER FILING DATE: 1990-06-13
EARLIER APPLICATION NUMBER: DE P40 18 988.0
NUMBER OF SEQ ID NOS: 16
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 11
LENGTH: 212
TYPE: PRT
ORGANISM: Borrelia burgdorferi
US-09-196-293-11

Query Match   34.0%;   Score 629;   DB 4;   Length 212;
Best Local Similarity 70.4%;   Pred. No. 1.8e-40;
Matches 138; Conservative 13; Mismatches 43; Indels 2; Gaps 2;

Qy   1 MACNNSGKDG-NTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAK 59

RESULT 13
US-08-209-603E-11
Sequence 11, Application US/0820603E
Patent No. 6248538
GENERAL INFORMATION:
~ APPLICANT: FUCHS, RENATE
~ APPLICANT: WILSKE, BETTINA
~ APPLICANT: PREAC-MURSIC, VERA
~ APPLICANT: MOTZ, MANFRED
~ APPLICANT: SOUTSCHEK, ERWIN
~ TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PROTEINS
~ TITLE OF INVENTION: FROM BORRELIA BURGDORFERI
~ NUMBER OF SEQUENCES: 15
~ CORRESPONDENCE ADDRESS:
~ ADDRESSEE: BROOKS HAIDT HAFFNER & DELAHUNTY
~ STREET: 99 PARK AVENUE
~ CITY: NEW YORK
~ STATE: NY
~ COUNTRY: USA
~ ZIP: 10016
~ COMPUTER READABLE FORM:
~ MEDIUM TYPE: 3.5" FLOPPY DISC
~ COMPUTER: AT&T - IBM COMPATIBLE
~ OPERATING SYSTEM: MS-DOS Version 6.2
~ SOFTWARE: ASCII
~ CURRENT APPLICATION DATA:
~ APPLICATION NUMBER: US/08/209,603E
~ FILING DATE: 10-MAR-1994
~ CLASSIFICATION: 436
~ PRIOR APPLICATION DATA:
~ APPLICATION NUMBER: PCT/EP90/02282
~ FILING DATE: 21-DEC-1990
~ APPLICATION NUMBER: US 07/862,535
~ FILING DATE: 19-JUN-1992
~ ATTORNEY/AGENT INFORMATION:
~ NAME: ROBINSON, WILLIAM R
~ REGISTRATION NUMBER: 27,224
~ REFERENCE/DOCKET NUMBER: LKR-9217-A
~ TELECOMMUNICATION INFORMATION:
~ TELEPHONE: (212) 697-3335
~ TELEFAX: (212) 557-5635
~ INFORMATION FOR SEQ ID NO: 11:
~ SEQUENCE CHARACTERISTICS:
~ LENGTH: 212
~ TYPE: AMINO ACID
~ TOPOLOGY: LINEAR
~ MOLECULE TYPE: PROTEIN
~ HYPOTHETICAL: N/A
~ ANTI-SENSE: N/A
~ FRAGMENT TYPE: N/A
~ ORIGINAL SOURCE:
~ ORGANISM: B. BURGDORFERI
~ IMMEDIATE SOURCE:
~ LIBRARY: DSM 5662
~ POSITION IN GENOME: N/A
~ FEATURE:

Qy   17 ISCNNSGKVGILTSTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETLVLSIDELAKK 76
Db   77 AIGQKIDNNNGLAALNNQNGSLLAGAYAISTLITEKLSKLKNLEELKTEIAKAKKCSEEF 136
Db 137 TNKLKSGHADLGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGLLKAAQVALTNS 196
Qy   60 AIGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKN-EGLKEKIDAARKCSETF 118
Db 197 VKELTSPVVAESPKKP 212
Qy 119 TNKLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANS 178
Qy 179 VKELTSPVVAESPKKP 194

~ IDENTIFICATION METHOD: amino acid analysis
~ PUBLICATION INFORMATION: N/A
~ ANTI-SENSE: N/A
~ FRAGMENT TYPE: INTERNAL
~ ORIGINAL SOURCE:
US-08-209-603E-11

Query Match   34.0%; Score 629; DB 4; Length 212;
Best Local Similarity 70.4%; Pred. No. 1.8e-40;
Matches 138; Conservative 13; Mismatches 43; Indels 2; Gaps 2;

Qy   1 MACNNSGKDG-NTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAK 59
Db   17 ISCNNSGKVGILTSTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETLVLSIDELAKK 76
Qy   60 AIGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKN-EGLKEKIDAARKCSETF 118
Db   77 AIGQKIDNNNGLAALNNQNGSLLAGAYAISTLITEKLSKLKNLEELKTEIAKAKKCSEEF 136
Qy 119 TNKLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANS 178
Db 137 TNKLKSGHADLGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGLLKAAQVALTNS 196
Qy 179 VKELTSPVVAESPKKP 194
Db 197 VKELTSPVVAESPKKP 212

RESULT 14
US-08-235-836C-34
Sequence 34, Application US/08235836C
Patent No. 6248562
GENERAL INFORMATION:
~ APPLICANT: Dunn, John J.
~ APPLICANT: Luft, Benjamin J.
~ TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising
~ TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor
~ NUMBER OF SEQUENCES: 144
~ CORRESPONDENCE ADDRESS:
~ ADDRESSEE: Brookhaven National Laboratory
~ STREET:
~ CITY: Upton
~ STATE: NY
~ COUNTRY: USA
~ ZIP: 11973
~ COMPUTER READABLE FORM:
~ MEDIUM TYPE: Floppy disk
~ COMPUTER: IBM PC compatible
~ OPERATING SYSTEM: PC-DOS/MS-DOS
~ SOFTWARE: Patentin Release #1.0, Version #1.25
~ CURRENT APPLICATION DATA:
~ APPLICATION NUMBER: US/08/235,836C
~ FILING DATE: 29-APR-1994
~ CLASSIFICATION: 435
~ PRIOR APPLICATION DATA:
~ APPLICATION NUMBER: US 08/148,191
~ FILING DATE: 01-11-93
~ ATTORNEY/AGENT INFORMATION:
~ NAME: Bogosian, Margaret C.
~ REGISTRATION NUMBER: 25,324
~ REFERENCE/DOCKET NUMBER: BNL93-28A
~ TELECOMMUNICATION INFORMATION:
~ TELEPHONE: (516) 282-7338
~ TELEFAX: (516) 282-3729
~ INFORMATION FOR SEQ ID NO: 34:
~ SEQUENCE CHARACTERISTICS:
~ LENGTH: 212 amino acids
~ TYPE: amino acid
~ TOPOLOGY: linear
~ MOLECULE TYPE: protein
US-08-235-836C-34

Query Match   33.8%; Score 626; DB 4; Length 212;
Best Local Similarity 69.4%; Pred. No. 3e-40;
Matches 136; Conservative 16; Mismatches 42; Indels 2; Gaps 2;

Qy   1 MACNNSGKDGNT-SANSADESVKGPNLTEINKKITDSNAVLLAVKEVEVLSSIDEIAAK 59
         ::!:!!!!!|!| :!!!!!:! !:!!!!!!!!!!!!!!!!!|!| :!:!!!!!!!
Db   17 ISCSNSGKGDSASTNPADSSAKGPNLTEISKKITDSNAFVLAVKEVEVLVLSIDELAKK 76

Qy   60 AIGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKN-EGLKEKIDAAKKCSETF 118
         |||:|| ||:|
Db   77 AIGQKIDNNNGLAALNNQGSLLAGAYAISTLITEKLSKLKNLEELKTEIAKAKKCSEEF 136

Qy 119 TNKLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANS 178
         ||||| | |||| | |||| ||||||| | ||| |:||| |||||:|| |||| |||
Db 137 TNKLKSGHADLGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGLLKAAQVALTNS 196

Qy 179 VKELTSPVVAESPKKP 194
         ||||||||||||||||
Db 197 VKELTSPVVAESPKKP 212

RESULT 15
US-08-235-836C-32
Sequence 32, Application US/08235836
Patent No. 6248562

GENERAL INFORMATION:
  APPLICANT: Dunn, John J.
  APPLICANT: Luft, Benjamin J.
  TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising
  TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor
  NUMBER OF SEQUENCES: 144
  CORRESPONDENCE ADDRESS:
    ADDRESSEE: Brookhaven National Laboratory
    STREET:
    CITY: Upton
    STATE: NY
    COUNTRY: USA
    ZIP: 11973
  COMPUTER READABLE FORM:
    MEDIUM TYPE: Floppy disk
    COMPUTER: IBM PC compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: Patentin Release #1.0, Version #1.25
  CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/08/235,836C
    FILING DATE: 29-APR-1994
    CLASSIFICATION: 435
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: US 08/148,191
    FILING DATE: 01-11-93
  ATTORNEY/AGENT INFORMATION:
    NAME: Bogosian, Margaret C.
    REGISTRATION NUMBER: 25,324
    REFERENCE/DOCKET NUMBER: BNL93-28A
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: (516) 282-7338
    TELEFAX: (516) 282-3729
INFORMATION FOR SEQ ID NO: 32:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 209 amino acids
    TYPE: amino acid
    TOPOLOGY: linear
    MOLECULE TYPE: protein

US-08-235-836C-32

Query Match   33.2%; Score 614.5; DB 4; Length 209;
Best Local Similarity 68.2%; Pred. No. 2.2e-39;
Matches 133; Conservative 20; Mismatches 39; Indels 3; Gaps 2;

Qy   1 MACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEVALLSSIDEIAAKA 60
         ::!||!!|| :!:.:! :!!! !||||||||||||||| :||||||||||| :||||||!| ||
Db   17 ISCNNSG--GDTASTNPDESAKGPNLTVISKKITDSNAFVLAVKEVEVEALLSSIDELANKA 74

Qy   61 IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKN-EGLKEKIDAAKKCSETFT 119
         |||:||||| | :|||||||||||||||||| :!:!:| || :!:!!!:!!!!|:!|!| ||
Db   75 IGKVIHQNGLNANAGQNGSLLAGAYAISTLITERLSKLKNSEELNKKIEEAKNHSEAFT 134

Qy 120 NKLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSV 179
         |:|| || || | |||:|||||||: |:||:||| :|| |:|||||:|| |||| |||||
Db 135 NRLKGSHAQLGVAAATDDHAKEAILKSNPTKDKGAKALKDLSESVESLAKAAQEALANSV 194

Qy 180 KELTSPVVAESPKKP 194
         |||: ||||||||||
Db 195 KELNPVVAESPKKP 209

Search completed: March 18, 2002, 09:55:33
Job time: 346 sec

OM protein - protein search, using sw model

Copyright (c) 1993 - 2000 Compugen Ltd.

GenCore version 4.5

Run on: March 18, 2002, 10:08:43 ; Search time 621.2 Seconds

168.507 Million cell updates/sec
(without alignments)

Title: US-09-596-746A-30
Perfect score: 1850
Sequence: 1 MACNNSGKDGNTSANSADES.........AVENLAKAAKEMLANSVKEL 377

Scoring table:
Gapop 10.0 , Gapext 0.5
BLOSUM62

Searched: 3148936 seqs, 27765734 residues

Post-processing: Minimum Match 0%
Maximum DB seq length: 0
Minimum DB seq length: 0
Maximum Match 100%
Listing first 45 summaries

Total number of hits satisfying chosen parameters: 3148936

Database :
Pending_Patents_AA_Main:*
1: /cgn2_6/ptodata/2/paa/PCTUS_COMB.pep:*
2: /cgn2_6/ptodata/2/paa/US06_COMB.pep:*
3: /cgn2_6/ptodata/2/paa/US07_COMB.pep:*
4: /cgn2_6/ptodata/2/paa/US80_COMB.pep:*
5: /cgn2_6/ptodata/2/paa/US81_COMB.pep:*
6: /cgn2_6/ptodata/2/paa/US82_COMB.pep:*
7: /cgn2_6/ptodata/2/paa/US83_COMB.pep:*
8: /cgn2_6/ptodata/2/paa/US84_COMB.pep:*
9: /cgn2_6/ptodata/2/paa/US85_COMB.pep:*
10: /cgn2_6/ptodata/2/paa/US86_COMB.pep:*
11: /cgn2_6/ptodata/2/paa/US87_COMB.pep:*
12: /cgn2_6/ptodata/2/paa/US88_COMB.pep:*
13: /cgn2_6/ptodata/2/paa/US89_COMB.pep:*
14: /cgn2_6/ptodata/2/paa/US90_COMB.pep:*
15: /cgn2_6/ptodata/2/paa/US91_COMB.pep:*
16: /cgn2_6/ptodata/2/paa/US92_COMB.pep:*
17: /cgn2_6/ptodata/2/paa/US93_COMB.pep:*
18: /cgn2_6/ptodata/2/paa/US94_COMB.pep:*
19: /cgn2_6/ptodata/2/paa/US95_COMB.pep:*
20: /cgn2_6/ptodata/2/paa/US96_COMB.pep:*
21: /cgn2_6/ptodata/2/paa/US97_COMB.pep:*
22: /cgn2_6/ptodata/2/paa/US98_COMB.pep:*
23: /cgn2_6/ptodata/2/paa/US99_COMB.pep:*
24: /cgn2_6/ptodata/2/paa/US60_COMB.pep:*

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1850	100.0	377	19	US-09-596-746A-30	Sequence 30, Appl
2	1850	100.0	377	19	US-09-596-746A-30	Sequence 30, Appl
3	1825	98.6	400	19	US-09-596-746A-82	Sequence 82, Appl
4	1825	98.6	401	19	US-09-596-746A-82	Sequence 82, Appl
5	1801.5	97.4	374	19	US-09-596-746A-26	Sequence 26, Appl
6	1796.5	97.1	373	19	US-09-596-746-26	Sequence 26, Appl
7	1793.5	96.9	397	19	US-09-596-746-62	Sequence 62, Appl
8	1793.5	96.9	398	19	US-09-596-746A-62	Sequence 62, Appl
9	1614.5	87.3	378	19	US-09-596-746-28	Sequence 28, Appl

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

10	1614.5	87.3	378	19	US-09-596-746-28	Sequence 28, Appl
11	1589.5	85.9	401	19	US-09-596-746A-80	Sequence 80, Appl
12	1589.5	85.9	402	19	US-09-596-746A-80	Sequence 80, Appl
13	1566	84.6	375	19	US-09-596-746-24	Sequence 24, Appl
14	1561	84.4	374	19	US-09-596-746A-24	Sequence 24, Appl
15	1558	84.2	398	19	US-09-596-746A-60	Sequence 60, Appl
16	1558	84.2	399	19	US-09-596-746A-60	Sequence 60, Appl
17	1524.5	82.4	370	19	US-09-596-746A-36	Sequence 36, Appl
18	1519.5	82.1	369	19	US-09-596-746-36	Sequence 36, Appl
19	1516.5	82.0	393	19	US-09-596-746-66	Sequence 66, Appl
20	1516.5	82.0	394	19	US-09-596-746A-66	Sequence 66, Appl
21	1485.5	80.3	368	19	US-09-596-746A-42	Sequence 42, Appl
22	1480.5	80.0	367	19	US-09-596-746-42	Sequence 42, Appl
23	1477.5	79.9	392	19	US-09-596-746-74	Sequence 74, Appl
24	1477.5	79.9	392	19	US-09-596-746A-74	Sequence 74, Appl
25	1472	79.6	385	19	US-09-596-746A-56	Sequence 56, Appl
26	1467	79.3	384	19	US-09-596-746-56	Sequence 56, Appl
27	1465	79.2	560	19	US-09-596-746-52	Sequence 52, Appl
28	1465	79.2	560	19	US-09-596-746A-52	Sequence 52, Appl
29	1464	79.1	408	19	US-09-596-746-78	Sequence 78, Appl
30	1464	79.1	409	19	US-09-596-746A-78	Sequence 78, Appl
31	1289	69.7	371	19	US-09-596-746A-34	Sequence 34, Appl
32	1284	69.4	370	19	US-09-596-746-34	Sequence 34, Appl
33	1281	69.2	394	19	US-09-596-746-64	Sequence 64, Appl
34	1281	69.2	395	19	US-09-596-746A-64	Sequence 64, Appl
35	1264.5	68.4	370	19	US-09-596-746-32	Sequence 32, Appl
36	1259.5	68.1	369	19	US-09-596-746A-32	Sequence 32, Appl
37	1256.5	67.9	393	19	US-09-596-746-68	Sequence 68, Appl
38	1256.5	67.9	394	19	US-09-596-746A-68	Sequence 68, Appl
39	1250	67.6	369	19	US-09-596-746-40	Sequence 40, Appl
40	1245	67.3	368	19	US-09-596-746-40	Sequence 40, Appl
41	1242	67.1	392	19	US-09-596-746-72	Sequence 72, Appl
42	1242	67.1	393	19	US-09-596-746A-72	Sequence 72, Appl
43	1225.5	66.2	368	19	US-09-596-746-38	Sequence 38, Appl
44	1225.5	66.2	368	19	US-09-596-746A-38	Sequence 38, Appl
45	1217.5	65.8	391	19	US-09-596-746-76	Sequence 76, Appl

ALIGNMENTS

RESULT 1
US-09-596-746-30

Sequence 30, Application US/0959674 6

GENERAL INFORMATION:
APPLICANT: Dattwyler, Raymond J.
APPLICANT: Seinost, Gerald
APPLICANT: Dykhuizen, Daniel
APPLICANT: Luft, Benjamin J.
APPLICANT: Maria J.C: Gomes-Solecki
TITLE OF INVENTION: Groups of Borrelia burgdorferi and
TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
FILE REFERENCE: 2631.1002-001
CURRENT APPLICATION NUMBER: US/09/596,746
CURRENT FILING DATE: 2000-06-16
PRIOR APPLICATION NUMBER: US 60/140,042
PRIOR FILING DATE: 1999-06-18
NUMBER OF SEQ ID NOS: 84
SOFTWARE: FastSEQ for Windows Version 4.0

SEQ ID NO 30
LENGTH: 377
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: OspC Chimera
US-09-596-746-30

Query Match   100.0%; Score 1850; DB 19; Length 377;
Best Local Similarity   100.0%; Pred. No. 3.4e-125;
Matches 377; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 MACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKA 60

RESULT 2
US-09-596-746A-30
Sequence 30, Application US/0959567446A
GENERAL INFORMATION:
APPLICANT: Dattwyler, Raymond J.
APPLICANT: Seinost, Gerald
APPLICANT: Dykhuizen, Danial
APPLICANT: Luft, Benjamin J.
APPLICANT: Maria J.C. Gomes-Solecki
TITLE OF INVENTION: Groups of Borrelia burgdorferi and
TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
FILE REFERENCE: 2631.1002-001
CURRENT APPLICATION NUMBER: US/09/596,746A
CURRENT FILING DATE: 2000-06-19
PRIOR APPLICATION NUMBER: US 60/140,042
PRIOR FILING DATE: 1999-06-18
NUMBER OF SEQ ID NOS: 84
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 30
LENGTH: 377
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: OspC Chimera
US-09-596-746A-30

Query Match   100.0%; Score 1850; DB 19; Length 377;
Best Local Similarity 100.0%; Pred. No. 3.4e-125;
Matches 377; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 MACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKA 60

Qy   61 IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN 120

Qy   121 KLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 KLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 180

Qy   181 ELTSPVVAESPKKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEI 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 ELTSPVVAESPKKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEI 240

Qy   241 ETLLASIDELATKAIGKKIQQNGGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLK 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 ETLLASIDELATKAIGKKIQQNGGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLK 300

Qy   301 EKIENAKKCSEDFTKKLEGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVE 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 EKIENAKKCSEDFTKKLEGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVE 360

Qy   361 NLAKAAKEMLANSVKEL 377
          |||||||||||||||||
Db   361 NLAKAAKEMLANSVKEL 377

RESULT 3
US-09-596-746-82
Sequence 82, Application US/0959567446
GENERAL INFORMATION:
APPLICANT: Dattwyler, Raymond J.
APPLICANT: Seinost, Gerald
APPLICANT: Dykhuizen, Danial
APPLICANT: Luft, Benjamin J.
APPLICANT: Maria J.C. Gomes-Solecki
TITLE OF INVENTION: Groups of Borrelia burgdorferi and
TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
FILE REFERENCE: 2631.1002-001
CURRENT APPLICATION NUMBER: US/09/596,746
CURRENT FILING DATE: 2000-06-16
PRIOR APPLICATION NUMBER: US 60/140,042
PRIOR FILING DATE: 1999-06-18
NUMBER OF SEQ ID NOS: 84
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 82
LENGTH: 400
TYPE: PRT
ORGANISM: OspC Chimera
US-09-596-746-82

Query Match   98.6%; Score 1825; DB 19; Length 400;
Best Local Similarity 98.9%; Pred. No. 2.4e-123;
Matches 372; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy   2 ACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKAI 61
          :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   25 SCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAI 84

Qy   62 GKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNK 121
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   85 GKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNK 144

Qy   122 LKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 181
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   145 LKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 204

Qy   182 LTSPVVAESPKKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIE 241
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   205 LTSPVVAESPKKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIE 264

Qy   242 TLLASIDELATKAIGKKIQQNGGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKE 301
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   265 TLLASIDELATKAIGKKIQQNGGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKE 324

Qy   302 KIENAKKCSEDFTKKLEGEHAQLGIENVTDENAKKAILITDAARDKGAAELEKLFKAVEN 361
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   325 KIENAKKCSEDFTKKLEGEHAQLGIENVTDENAKKAILITDAARDKGAAELEKLFKAVEN 384

Qy   362 LAKAAKEMLANSVKEL 377
          ||||||||||||||||
Db   385 LAKAAKEMLANSVKEL 400

RESULT 4
US-09-596-746A-82
Sequence 82, Application US/0959746A
GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seinost, Gerald
; APPLICANT: Dykhuizen, Danial
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746A
; CURRENT FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 82
;   LENGTH: 401
;   TYPE: PRT
;   ORGANISM: ospc Chimera
US-09-596-746A-82

Query Match   98.6%;   Score 1825; DB 19; Length 401;
Best Local Similarity   98.9%;   Pred. No. 2.4e-123;
Matches   372; Conservative   2; Mismatches   2; Indels   0; Gaps   0;

Qy   2 ACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKAI   61
           ::||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   26 SCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAI   85

Qy   62 GKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNK 121
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   86 GKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNK 145

Qy   122 LKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 181
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   146 LKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 205

Qy   182 LTSPVVAESPKRPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIE 241
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   206 LTSPVVAESPKRPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIE 265

Qy   242 TLLASIDELATKAIGKKIQQNGGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKE 301
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   266 TLLASIDELATKAIGKKIQQNGGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKE 325

Qy   302 KIENAKKCSEDFTKKLEGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVEN 361
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   326 KIENAKKCSEDFTKKLEGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVEN 385

Qy   362 LAKAAKEMLANSVKEL 377
           ||||||||||||||||
Db   386 LAKAAKEMLANSVKEL 401

RESULT 5
US-09-596-746A-26
Sequence 26, Application US/0959746A
GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seinost, Gerald
; APPLICANT: Dykhuizen, Danial
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746A
; CURRENT FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18

Query Match   97.4%;   Score 1801.5; DB 19; Length 374;
Best Local Similarity   98.1%;   Pred. No. 1.1e-121;
Matches   370; Conservative   2; Mismatches   2; Indels   3; Gaps   1;

Qy   1 MACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKA   60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA   60

Qy   61 IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN 120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 IGKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN 120

Qy   121 KLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 180
           ||||||||||||||||||||||||||:|||||||||||||||||||||||||||||||||
Db   121 KLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 180

Qy   181 ELTSPVVAESPKRPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEI 240
           ||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db   181 ELTSPVVAES---PAMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEI 237

Qy   241 ETLLASIDELATKAIGKKIQQNGGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLK 300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   238 ETLLASIDELATKAIGKKIQQNGGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLK 297

Qy   301 EKIENAKKCSEDFTKKLEGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVE 360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   298 EKIENAKKCSEDFTKKLEGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVE 357

Qy   361 NLAKAAKEMLANSVKEL 377
           ||||||||||||||||
Db   358 NLAKAAKEMLANSVKEL 374

RESULT 6
US-09-596-746A-26
Sequence 26, Application US/0959746A
GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seinost, Gerald
; APPLICANT: Dykhuizen, Danial
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746
; CURRENT FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 26
;   LENGTH: 373
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: OspC Chimera
US-09-596-746A-26

Query Match   97.1%;   Score 1796.5; DB 19; Length 373;

```
Best Local Similarity  98.1%;  Pred. No. 2.4e-121;
Matches 369; Conservative 2; Mismatches 2; Indels 3; Gaps 1;

Qy    2 ACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKAI  61
         |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db    1 ACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAI  60

Qy   62 GKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNK 121
         ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db   61 GKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNK 120

Qy  122 LKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKRAAKEMLANSVKE 181
         |||||||||||||||||||||||| ||||||||||||||||||||||| |||||||||||
Db  121 LKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 180

Qy  182 LTSPVVAESPKKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIE 241
         ||||||||    :|||||||||||||||||||||||||||||||||||||||||||||||
Db  181 LTSPVVAES---PAMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIE 237

Qy  242 TLLASIDELATKAIGKKIQONGGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKE 301
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  238 TLLASIDELATKAIGKKIQONGGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKE 297

Qy  302 KIENAKKCSEDFTKKLEGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVEN 361
         |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db  298 KIENAKKCSEDFTKKLEGEHAQLGIENVTDENAKRAILITDAAKDKGAAELEKLFKAVEN 357

Qy  362 LAKAAKEMLANSVKEL 377
         ||||||||||||||||
Db  358 LAKAAKEMLANSVKEL 373


RESULT 7
US-09-596-746-62
Sequence 62, Application US/0956746
GENERAL INFORMATION:
  APPLICANT: Dattwyler, Raymond J.
  APPLICANT: Seinost, Gerald
  APPLICANT: Dykhuizen, Daniel
  APPLICANT: Luft, Benjamin J.
  APPLICANT: Maria J.C. Gomes-Solecki
  TITLE OF INVENTION: Groups of Borrelia burgdorferi and
  TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
  FILE REFERENCE: 2631.1002-001
  CURRENT APPLICATION NUMBER: US/09,596,746
  CURRENT FILING DATE: 2000-06-16
  PRIOR APPLICATION NUMBER: US 60/140,042
  PRIOR FILING DATE: 1999-06-18
  SOFTWARE: FastSEQ for Windows Version 4.0
  SEQ ID NO 62
  LENGTH: 397
  TYPE: PRT
  ORGANISM: ospc Chimera
US-09-596-746-62

Query Match     96.9%;  Score 1793.5;  DB 19;  Length 397;
Best Local Similarity  97.9%;  Pred. No. 4.4e-121;
Matches 368; Conservative 3; Mismatches 2; Indels 3; Gaps 1;

Qy    2 ACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKAI  61
         :|||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db   25 SCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAI  84

Qy   62 GKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNK 121
         ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db   85 GKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNK 144

Qy  122 LKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKRAAKEMLANSVKE 181
         |||||||||||||||||||||||| ||||||||||||||||||||||| |||||||||||
Db  145 LKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 204

Qy  182 LTSPVVAESPKKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIE 241
         ||||||||    :|||||||||||||||||||||||||||||||||||||||||||||||
Db  205 LTSPVVAES---PAMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIE 261

Qy  242 TLLASIDELATKAIGKKIQONGGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKE 301
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  262 TLLASIDELATKAIGKKIQONGGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKE 321

Qy  302 KIENAKKCSEDFTKKLEGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVEN 361
         |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db  322 KIENAKKCSEDFTKKLEGEHAQLGIENVTDENAKRAILITDAAKDKGAAELEKLFKAVEN 381

Qy  362 LAKAAKEMLANSVKEL 377
         ||||||||||||||||
Db  382 LAKAAKEMLANSVKEL 397
```

```
RESULT 8
US-09-596-746A-62
Sequence 62, Application US/0956746A
GENERAL INFORMATION:
  APPLICANT: Dattwyler, Raymond J.
  APPLICANT: Seinost, Gerald
  APPLICANT: Dykhuizen, Daniel
  APPLICANT: Luft, Benjamin J.
  APPLICANT: Maria J.C. Gomes-Solecki
  TITLE OF INVENTION: Groups of Borrelia burgdorferi and
  TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
  FILE REFERENCE: 2631.1002-001
  CURRENT APPLICATION NUMBER: US/09,596,746A
  CURRENT FILING DATE: 2000-06-19
  PRIOR APPLICATION NUMBER: US 60/140,042
  PRIOR FILING DATE: 1999-06-18
  SOFTWARE: FastSEQ for Windows Version 4.0
  SEQ ID NO 62
  LENGTH: 398
  TYPE: PRT
  ORGANISM: ospc Chimera
US-09-596-746A-62

Query Match     96.9%;  Score 1793.5;  DB 19;  Length 398;
Best Local Similarity  97.9%;  Pred. No. 4.4e-121;
Matches 368; Conservative 3; Mismatches 2; Indels 3; Gaps 4;

Qy    2 ACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKAI  61
         :|||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db   26 SCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAI  85

Qy   62 GKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNK 121
         ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db   86 GKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNK 145

Qy  122 LKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKRAAKEMLANSVKE 181
         |||||||||||||||||||||||| ||||||||||||||||||||||| |||||||||||
Db  146 LKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 205

Qy  182 LTSPVVAESPKKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIE 241
         ||||||||    :|||||||||||||||||||||||||||||||||||||||||||||||
Db  206 LTSPVVAES---PAMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIE 262

Qy  242 TLLASIDELATKAIGKKIQONGGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKE 301
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  263 TLLASIDELATKAIGKKIQONGGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKE 322

Qy  302 KIENAKKCSEDFTKKLEGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVEN 361
         |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db  323 KIENAKKCSEDFTKKLEGEHAQLGIENVTDENAKRAILITDAAKDKGAAELEKLFKAVEN 382

Qy  362 LAKAAKEMLANSVKEL 377
         ||||||||||||||||
Db  383 LAKAAKEMLANSVKEL 398
```

RESULT 9
US-09-596-746-28
Sequence 28, Application US/0596746
GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seinost, Gerald
; APPLICANT: Dykhuizen, Danial
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746
; CURRENT FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 28
; LENGTH: 378
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: OspC Chimera
US-09-596-746-28

Query Match 87.3%; Score 1614.5; DB 19; Length 378;
Best Local Similarity 87.8%; Pred. No. 3.6e-108;
Matches 331; Conservative 20; Mismatches 25; Indels 1; Gaps 1;

Qy   1 MACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKA 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKA 60

Qy   61 IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN 120

Qy 121 KLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 KLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 180

Qy 181 ELTSPVVAESPKKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEV 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||| :
Db 181 ELTSPVVAESPKKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEI 240

Qy 241 ETLLTSIDELA-KAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELK 299
         : ||| :|||: || :: :: | ||::|:||::|| ::: || :::|::||:: || :||
Db 241 EKIENAKKCSEDFTKKLEGEHAOLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVE 300

Qy 301 EKIENAKKCSEDFTKKLEGEHAOLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVE 360
         |:|| ||||||||||||||||||||| |::|| :: | || ||||||||||||||:: :
Db 300 AEIEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVK 359

Qy 361 NLAKAAKEMLANSVKEL 377
         |:|:||||||||||||
Db 360 NLSKAAKEMLTNSVKEL 376

RESULT 10
US-09-596-746A-28
Sequence 28, Application US/0596746A
GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seinost, Gerald
; APPLICANT: Dykhuizen, Danial
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans

FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746A
; CURRENT FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 28
; LENGTH: 378
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: OspC Chimera
US-09-596-746A-28

Query Match 87.3%; Score 1614.5; DB 19; Length 378;
Best Local Similarity 87.8%; Pred. No. 3.6e-108;
Matches 331; Conservative 20; Mismatches 25; Indels 1; Gaps 1;

Qy   1 MACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKA 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKA 60

Qy   61 IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN 120

Qy 121 KLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 KLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 180

Qy 181 ELTSPVVAESPKKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEV 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||| :
Db 181 ELTSPVVAESPKKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEI 240

Qy 241 ETLLTSIDELA-KAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELK 299
         : ||| :|||: || :: :: | ||::|:||::|| ::: || :::|::||:: || :||
Db 241 EKIENAKKCSEDFTKKLEGEHAOLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVE 300

Qy 301 EKIENAKKCSEDFTKKLEGEHAOLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVE 360
         |:|| ||||||||||||||||||||| |::|| :: | || ||||||||||||||:: :
Db 300 AEIEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVK 359

Qy 361 NLAKAAKEMLANSVKEL 377
         |:|:||||||||||||
Db 360 NLSKAAKEMLTNSVKEL 376

RESULT 11
US-09-596-746-80
Sequence 80, Application US/0596746
GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seinost, Gerald
; APPLICANT: Dykhuizen, Danial
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746
; CURRENT FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 80
; LENGTH: 401
; TYPE: PRT
; ORGANISM: OspC Chimera
US-09-596-746-80

Query Match   85.9%;   Score 1589.5; DB 19; Length 401;
Best Local Similarity 86.7%; Pred. No. 2.5e-106;
Matches 326; Conservative 22; Mismatches 27; Indels 1; Gaps 1;

RESULT 12
US-09-596-746A-80
; Sequence 80, Application US/0959746A
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seinost, Gerald
; APPLICANT: Dykhuizen, Danial
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746A
; CURRENT FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 80
; LENGTH: 402
; TYPE: PRT
; ORGANISM: ospc Chimera
US-09-596-746A-80

Qy   2 ACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKAI 61
         ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db   25 SCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAI 84

Qy   62 GKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNK 121
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   85 GKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNK 144

Qy 122 LKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 181
         ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db 145 LKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 204

Qy 182 LTSPVVAESPKKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVE 241
         ||||||||||||||||||||||||||||||||||||||:|||||||||||||:|:||||
Db 205 LTSPVVAESPKKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVE 264

Qy 242 TLLASIDELATKAIGKKIQONGGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKE 301
         |||||||||||| ||||||:||||||||||||||||||||||:|| :||:| :|:| ||
Db 265 TLLTSIDELA-KAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKISAIKDSGELKA 323

Qy 302 KIENAKKCSEDFTKKLEGEHAOLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVEN 361
         :|||||||||||||||||:::|||||||||||||||||||| |||| ||||||:|:|:|
Db 324 EIEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKN 383

Qy 362 LAKAAKEMLANSVKEL 377
         :|||||||||||||||
Db 384 LSKAAKEMLTNSVKEL 399

Query Match   85.9%;   Score 1589.5; DB 19; Length 402;
Best Local Similarity 86.7%; Pred. No. 2.5e-106;
Matches 326; Conservative 22; Mismatches 27; Indels 1; Gaps 1;

Qy   2 ACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKAI 61
         ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db   26 SCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAI 85

Qy   62 GKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNK 121
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   86 GKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNK 145

Qy 122 LKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 181

RESULT 13
US-09-596-746A-24
; Sequence 24, Application US/0959746A
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seinost, Gerald
; APPLICANT: Dykhuizen, Danial
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746A
; CURRENT FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: OspC Chimera
US-09-596-746A-24

Db 146 LKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 205

Qy 182 LTSPVVAESPKKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIE 241
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||:|:|
Db 206 LTSPVVAESPKKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVE 265

Qy 242 TLLASIDELATKAIGKKIQONGGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKE 301
         |||||||||| ||||||:||||||||||||||||||||||||:||:|| : ||:|| ||
Db 266 TLLTSIDELA-KAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKISAIKDSGELKA 323

Qy 302 KIENAKKCSEDFTKKLEGEHAOLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVE 360
         :|||||||||||||||||:::||||||||||||||||||| ||||| ||||||:|:|:
Db 325 EIEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVK 384

Qy 362 LAKAAKEMLANSVKEL 377
         :|||||||||||||||
Db 385 LSKAAKEMLTNSVKEL 400

Query Match   84.6%;   Score 1566; DB 19; Length 375;
Best Local Similarity 85.9%; Pred. No. 1.1e-104;
Matches 324; Conservative 22; Mismatches 27; Indels 27; Gaps 2;

Qy   1 MACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKA 60
         |||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db   1 MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA 60

Qy   61 IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 IGKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN 120

Qy 121 KLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 180
         |||||||||||||||||||||||||:|||||||||||||||||||||||||||||||||
Db 121 KLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 180

Qy 181 ELTSPVVAESPKKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEI 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:
Db 181 ELTSPVVAES---PAMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEV 237

Qy 241 ETLLASIDELATKAIGKKIQONGGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLK 300
         |||||||||| ||||||:||||||||||||||||||||||||:||:||: ||:|: ||
Db 238 ETLLTSIDELA-KAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKISAIKDSGELK 296

Qy 301 EKIENAKKCSEDFTKKLEGEHAOLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVE 360
         :||||||||||||||||:::||||||||||||||||||||| |||| ||||||:|:|:
Db 297 AEIEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVK 356

```
Qy   361 NLAKAAKEMLANSVKEL 377
         |:|:|||||| |||||
Db   357 NLSKAAKEMLTNSVKEL 373

RESULT 14
US-09-596-746-24

Sequence 24, Application US/09596746

GENERAL INFORMATION:
  APPLICANT:  Dattwyler, Raymond J.
  APPLICANT:  Selnost, Gerald
  APPLICANT:  Dykhuizen, Danial
  APPLICANT:  Luft, Benjamin J.
  APPLICANT:  Maria J.C. Gomes-Solecki
  TITLE OF INVENTION:  Groups of Borrelia burgdorferi and
  TITLE OF INVENTION:  Borrelia afzelii That Cause Lyme Disease in Humans
  FILE REFERENCE:  2631.1002-001
  CURRENT APPLICATION NUMBER:  US/09/596,746
  CURRENT FILING DATE:  2000-06-16
  PRIOR APPLICATION NUMBER:  US 60/140,042
  PRIOR FILING DATE:  1999-06-18
  NUMBER OF SEQ ID NOS:  84
  SOFTWARE:  FastSEQ for Windows Version 4.0
  SEQ ID NO 24
  LENGTH:  374
  TYPE:  PRT
  ORGANISM:  Artificial Sequence
  FEATURE:
  OTHER INFORMATION:  OspC Chimera
US-09-596-746-24

Query Match   84.4%;  Score 1561;  DB 19;  Length 374;
Best Local Similarity  85.9%;  Pred. No. 2.6e-104;
Matches 323;  Conservative 22;  Mismatches 27;  Indels  4;  Gaps  2;

Qy     2 ACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKAI 61
         |||||||||||||||||||||||||||::||||||||||||||||||||||||||||||
Db     1 ACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAI 60

Qy    62 GKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNK 121
         ||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db    61 GKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNK 120

Qy   122 LKEKHTDLGKEGYTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSYKE 181
         ||||||||||||:||||||||||:|||||||||||||||||||||||||||||||||:||
Db   121 LKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 180

Qy   182 LTSPVVAESPKKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIE 241
         |||||||||   ||:|||||||||||||||||||||||||||||||||||||||||||:|
Db   181 LTSPVVAES---PAMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVE 237

Qy   242 TLLASIDELATKAIGKKIQONGGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKE 301
         |||:||||| |||||||:::: : ::: ||||| :|||| ||:|||.:|: |: :| ||
Db   238 TLLTSIDELA-KAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKA 296

Qy   302 KIENAKKCSEDFTKKLEGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVEN 361
         :|:.|||||:.||.||:||:: ||:|||:||||||||: :|: ||||| :|||||: |:|
Db   297 EIEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKN 356

Qy   362 LAKAAKEMLANSVKEL 377
         |:|||||||:||||||
Db   357 LSKAAKEMLTNSVKEL 372
```

```
  APPLICANT:  Luft, Benjamin J.
  APPLICANT:  Maria J.C. Gomes-Solecki
  TITLE OF INVENTION:  Groups of Borrelia burgdorferi and
  TITLE OF INVENTION:  Borrelia afzelii That Cause Lyme Disease in Humans
  FILE REFERENCE:  2631.1002-001
  CURRENT APPLICATION NUMBER:  US/09/596,746
  CURRENT FILING DATE:  2000-06-16
  PRIOR APPLICATION NUMBER:  US 60/140,042
  PRIOR FILING DATE:  1999-06-18
  NUMBER OF SEQ ID NOS:  84
  SOFTWARE:  FastSEQ for Windows Version 4.0
  SEQ ID NO 60
  LENGTH:  398
  TYPE:  PRT
  ORGANISM:  OspC Chimera
US-09-596-746-60

Query Match   84.2%;  Score 1558;  DB 19;  Length 398;
Best Local Similarity  85.6%;  Pred. No. 4.7e-104;
Matches 322;  Conservative 23;  Mismatches 27;  Indels  4;  Gaps  2;

Qy     2 ACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKAI 61
         :||||||||||||||||||||||||||::||||||||||||||||||||||||||||||
Db    25 SCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAI 84

Qy    62 GKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNK 121
         ||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db    85 GKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNK 144

Qy   122 LKEKHTDLGKEGYTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSYKE 181
         ||||||||||||:||||||||||:|||||||||||||||||||||||||||||||||:||
Db   145 LKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 204

Qy   182 LTSPVVAESPKKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVE 241
         |||||||||   ||:|||||||||||||||||||||||||||||||||||||||||||:|
Db   205 LTSPVVAES---PAMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVE 261

Qy   242 TLLASIDELATKAIGKKIQONGGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKE 301
         |||:||||| |||||||:::: : ::: ||||| :|||| ||:|||.:|: |: :| ||
Db   262 TLLTSIDELA-KAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKA 320

Qy   302 KIENAKKCSEDFTKKLEGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVEN 361
         :|:.|||||:.||.||:||:: ||:|||:||||||||: :|: ||||| :|||||: |:|
Db   321 EIEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKN 380

Qy   362 LAKAAKEMLANSVKEL 377
         |:|||||||:||||||
Db   381 LSKAAKEMLTNSVKEL 396
```

```
RESULT 15
US-09-596-746-60

Sequence 60, Application US/09596746

GENERAL INFORMATION:
  APPLICANT:  Dattwyler, Raymond J.
  APPLICANT:  Selnost, Gerald
  APPLICANT:  Dykhuizen, Danial
```

```
Search completed: March 18, 2002, 10:08:44
Job time: 972 sec
```

OM protein - protein search, using sw model

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

Run on: March 18, 2002, 09:58:19 ; Search time 81.02 Seconds
    (without alignments)
    297.254 Million cell updates/sec

Title: US-09-596-746A-30
Perfect score: 1850
Sequence: 1 MACNNSGKDGNTSANSADES.........AVENLAKAAKEMLANSVKEL 377

Scoring table: BLOSUM62
     Gapop 10.0 , Gapext 0.5

Searched: 303921 seqs, 63882009 residues

Total number of hits satisfying chosen parameters: 303921

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
     Maximum Match 100%
     Listing first 45 summaries

Database :
Pending_Patents_AA_New:*
1: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pep:*
2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep1:*
7: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep:*
8: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	%	Query Match Length	DB	ID	Description
1	692.5	37.4	211	6	US-09-974-992-5	Sequence 5, Appli
2	631	34.1	212	6	US-09-974-992-7	Sequence 7, Appli
3	596.5	32.2	207	6	US-09-974-992-3	Sequence 3, Appli
4	172.5	9.3	1014	6	US-09-708-427-19883	Sequence 19883, A
5	172.5	9.3	1018	6	US-09-708-427-19882	Sequence 19882, A
6	172.5	9.3	1269	6	US-09-708-427-19881	Sequence 19881, A
7	168	9.1	630	6	US-09-742-096-5	Sequence 5, Appli
8	160	8.6	1786	6	US-09-742-096-3	Sequence 3, Appli
9	159.5	8.5	1639	6	US-09-614-150-6213	Sequence 6213, Ap
10	156.5	8.5	805	6	US-09-708-427-3730	Sequence 3730, Ap
11	156.5	8.5	841	6	US-09-708-427-3729	Sequence 3729, Ap
12	155.5	8.4	1038	6	US-09-708-427-32624	Sequence 32624, A
13	155.5	8.4	1043	6	US-09-708-427-32622	Sequence 32622, A
14	155.5	8.4	1042	6	US-09-708-427-32623	Sequence 32623, A.
15	153	8.3	1881	6	US-10-032-585-7646	Sequence 7646, Ap
16	153	8.3	1881	7	US-10-072-851-15590	Sequence 15590, A
17	152.5	8.2	2437	6	US-09-815-242-5834	Sequence 5834, Ap
18	152.5	8.2	2437	6	US-10-072-851-5834	Sequence 5834, Ap
19	152.5	8.2	6281	6	US-09-815-242-12996	Sequence 12996, A
20	152.5	8.2	6281	7	US-10-072-851-12996	Sequence 12996, A
21	152	8.2	2086	6	US-09-815-242-5639	Sequence 5639, Ap
22	152	8.2	2086	7	US-10-072-851-5639	Sequence 5639, Ap
23	152	8.2	2086	6	US-09-815-242-12610	Sequence 12610, A
24	152	8.2	5795	6	US-10-072-851-12610	Sequence 12610, A
25	151.5	8.2	2056	6	US-09-614-150-4824	Sequence 4824, Ap
26	150.5	8.1	2434	6	US-09-815-242-5835	Sequence 5835, Ap
27	150.5	8.1	2434	7	US-10-072-851-5835	Sequence 5835, Ap
28	150	8.1	837	6	US-09-815-242-5883	Sequence 5883, Ap
29	150	8.1	837	7	US-10-072-851-5883	Sequence 5883, Ap
30	150	8.1	875	6	US-09-815-242-13080	Sequence 13080, A
31	150	8.1	875	7	US-10-072-851-13080	Sequence 13080, A
32	148.5	8.0	730	6	US-09-708-427-3731	Sequence 3731, Ap
33	148.5	8.0	1313	6	US-09-708-427-15044	Sequence 15044, A
34	148	8.0	3070	6	US-09-961-403-7	Sequence 7, Appl1
35	148	8.0	3110	6	US-09-561-709B-7	Sequence 7, Appl1
36	147.5	8.0	738	6	US-09-708-427-21266	Sequence 21266, A
37	147.5	8.0	763	6	US-09-708-427-21267	Sequence 21267, A
38	147.5	7.9	1304	6	US-09-708-427-15045	Sequence 15045, A
39	147.5	7.9	1690	6	US-09-614-150-10224	Sequence 10224, A
40	145.5	7.7	1690	6	US-09-614-150-10311	Sequence 10311, A
41	145.5	7.7	1144	6	US-09-708-427-15046	Sequence 15046, A
42	142.5	7.7	1489	6	US-09-708-427-6636	Sequence 6636, Ap
43	142	7.7	1725	7	US-10-037-182-12	Sequence 12, Appl
44	141.5	7.6	1725	7	US-10-037-182-12	Sequence 12, Appl
45	141.5	7.6	1786	5	US-10-037-182-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1
US-09-974-992-5
Sequence 5, Application US/09974992
GENERAL INFORMATION:
APPLICANT: Mathiesen, Marianne J.
APPLICANT: Theisen, Michael
APPLICANT: Holm, Arne
APPLICANT: Ostergaard, Soren
TITLE OF INVENTION: Novel OspC-derived peptide fragments
FILE REFERENCE: 459-666P
CURRENT APPLICATION NUMBER: US/09,974,992
CURRENT FILING DATE: 2001-10-10
PRIOR APPLICATION NUMBER: US/09,180,089
PRIOR FILING DATE: 1999-05-13
PRIOR APPLICATION NUMBER: PCT/DK97/00203
PRIOR FILING DATE: 1997-05-02
NUMBER OF SEQ ID NOS: 40
SOFTWARE: Patentin Ver. 2.1

SEQ ID NO 5
LENGTH: 211
TYPE: PRT
ORGANISM: Borrelia burgdorferi
US-09-974-992-5

Query Match 37.4%;   Score 692.5;   DB 6;   Length 211;
Best Local Similarity   77.0%;   Pred. No. 3.5e-43;
Matches 151;   Conservative 12;   Mismatches 30;   Indels 3;   Gaps 3;

```
Qy    1 MACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKA 60
         :|||||||||||||||||||||||| ||||||||||||||  |||||||||||| ||||
Db   17 ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEL-AKA 75

Qy   61 IGKKIHONNGLDTENNINGSLLAGAYAYAISTLIKQKLFDGLK-NEGLKEKIDAAKKCSETFT 119
         ||||  ||| | ||||| |||||| |||||| ||| |  :|||||| :|
Db   76 IGKKIKNDGSLGDEANHNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEEFS 135

Qy  120 NKLKEKHTDLGKEGVTDADAKEAILKANGT-KTKGAEELGKLFESVEVLSKAAKEMLANS 178
         |||: ||| || | |||||| ||:|||||  || || || ||||||||  |||||||||
Db  136 TKLKDNHAQLGIQGVTDENAKRAILKANAAGKDKDGVEELEKLSGSLESLSKAAKEMLANS 195

Qy  179 VKELTSPVVAESPKKP 194
         |||||||| |||||||
Db  196 VKELTSPVVVESPKKP 211
```

RESULT 2
US-09-974-992-7

; Sequence 7, Application US/09974992
; GENERAL INFORMATION:
;   APPLICANT: Mathiesen, Marianne J.
;   APPLICANT: Theisen, Michael
;   APPLICANT: Holm, Arne
;   APPLICANT: Ostergaard, Soren
;   TITLE OF INVENTION: Novel OspC-derived peptide fragments
;   FILE REFERENCE: 459-666P
;   CURRENT APPLICATION NUMBER: US/09/974,992
;   CURRENT FILING DATE: 2001-10-10
;   PRIOR APPLICATION NUMBER: 09/180,089
;   PRIOR FILING DATE: 1999-05-13
;   PRIOR APPLICATION NUMBER: PCT/DK97/00203
;   PRIOR FILING DATE: 1997-05-02
;   NUMBER OF SEQ ID NOS: 40
;   SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
;   LENGTH: 212
;   TYPE: PRT
;   ORGANISM: Borrelia afzelii
US-09-974-992-7

Query Match   34.1%;   Score 631;   DB 6;   Length 212;
Best Local Similarity   69.9%;   Pred. No. 1e-38;
Matches 137; Conservative 15; Mismatches 42;   Indels   2;   Gaps   2;

Qy   1 MACNNSGKDGNT-SANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAK 59
          :::|||||| |:::|| || ||||||||||||||||||||||||||||||||:::| ||
Db   17 ISCNNSG--GDSASTNPDESAKGPNLTEISKKIITDSNAFLLAVKEVEALLSSIDEL-SKA 73

Qy   60 AIGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKN-EGLKEKIDAAKKCSEFF 118
          | ||||||||||||||||||||||||||||||||||| ||| ||||||| ||||||| |
Db   77 AIGQKIDNNNGLAALNNONGSLLAGAYAISTLIITEKLSKLKNLEELKTEIAKAKKCSEEF 136

Qy 119 TNKLKEKHTDLGKEGYTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANS 178
          ||||| || |:| |||||||||||||||||||| :|||||||||||||||||||||||
Db 137 TNKLKSGHADLGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGLLKAAQVALTNS 196

Qy 179 VKELTSPVVAESPKKP 194
          |||:||||||||||||
Db 197 VKELTSPVVAESPKKP 212

RESULT 3
Sequence 3, Application US/09974992
; GENERAL INFORMATION:
;   APPLICANT: Mathiesen, Marianne J.
;   APPLICANT: Theisen, Michael
;   APPLICANT: Holm, Arne
;   APPLICANT: Ostergaard, Soren
;   TITLE OF INVENTION: Novel OspC-derived peptide fragments
;   FILE REFERENCE: 459-666P
;   CURRENT APPLICATION NUMBER: US/09/974,992
;   CURRENT FILING DATE: 2001-10-10
;   PRIOR APPLICATION NUMBER: 09/180,089
;   PRIOR FILING DATE: 1999-05-13
;   PRIOR APPLICATION NUMBER: PCT/DK97/00203
;   PRIOR FILING DATE: 1997-05-02
;   NUMBER OF SEQ ID NOS: 40
;   SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
;   LENGTH: 207
;   TYPE: PRT
;   ORGANISM: Borrelia garinii
US-09-974-992-3

Query Match   32.2%;   Score 596.5;   DB 6;   Length 207;
Best Local Similarity   66.0%;   Pred. No. 3.2e-36;
Matches 128; Conservative 25; Mismatches 38;   Indels   3;   Gaps   2;

RESULT 4
US-09-708-427-19883
; Sequence 19883, Application US/09708427
; GENERAL INFORMATION:
;   APPLICANT: N. ALEXANDROV et al.
;   TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
;   TITLE OF INVENTION: THEREBY
;   FILE REFERENCE: 2750-1243P
;   CURRENT APPLICATION NUMBER: US/09/708,427
;   CURRENT FILING DATE: 2000-11-09
;   NUMBER OF SEQ ID NOS: 85364
;   SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19883
;   LENGTH: 1014
;   TYPE: PRT
;   ORGANISM: Arabidopsis thaliana
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 1..1014
;   OTHER INFORMATION: THEREBY
;   NAME/KEY: misc_feature
;   LOCATION: 1..1014
;   OTHER INFORMATION: Xaa is any amino acid
;   NAME/KEY: misc_feature
;   LOCATION: 1..1014
;   OTHER INFORMATION: Ceres Seq. ID 1836847
US-09-708-427-19883

Query Match   9.3%;   Score 172.5;   DB 6;   Length 1014;
Best Local Similarity   23.2%;   Pred. No. 0.00018;
Matches 113; Conservative 78; Mismatches 161;   Indels 135;   Gaps 22;

Qy   8 KDGNTSANSADE---------------SVKGPNL---------TEINKKITDS-NAVLLAVKE 45
          ||| :|||| ||| ||| |||||||| || | |
Db 401 KHGETEADSKGYLGQVAELQSTLEAFQVKSSSLEAALNIATENEKELTENLNAVTSEKKK 460

Qy   46 VEALLSSIDEIAAK-----------AIGKKIHQNNGLDTENNHNGSLLAGAYAISTL--IKQ 94
          :|| :| || | |||||||||||||||||||||||||||| ||
Db 461 LEA---TVDEYSVKIESENLLESIRNELNVTQG-KLESIENDLKAAGLQESEVMEKLKS 516

Qy   95 KLDGLKNEGLKEKIDAA----------------------KKCSEFFTNK-------LKEKH 126
          |||| | | |||||| ::| ||| | ||||
Db 517 AEESLEQKG--REIDEATTKRMELFALHQSLSIDSEHRLQKAMEEFTSRDSEASSLIEKL 574

Qy 127 TDL-GKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKEL--- 182
          || | |: ||||||||||||:| || |: |:|| ||: | |
Db 575 RDLEGKIKSYEQLAFASGKSSSIKEKLEQTLGRLAAAAESVNEKLKQEFDQAQEKSLQSS 634

Qy 183 -TSPVVAESPKK--------------PSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAV 233
          |::|| || | |: |:|: || :: | |: | | :||||||| |
Db 635 SESELLAETNNQLKIKIQELEGLLGSGSVEKETALKRLEEAIE-----RFNQKETESSDL 689

Qy 234 VLAVKEIETLLASIDELATKAIG------------------KKIQQNG----GLAVEA 269
          | |:|: ||| | | | | ::| | | :|| |
Db 690 VEKLKTHENQIEEYKLAHEASGVADTRKVELEDALSKLKNLLESTIEELGARCQGLEKES 749

Qy 270 GHNGTL------LAGAYTISKLITQKLDGLK-NSEKLKEKIENAKKCSEDFTKKL--EG 319
          || |: :| : :|| || ||
Db 750 ...

Db 750 GDLAEVNLKLNLELANHGSEANELQTKLSALEAEKEQTANELEASKTTIEDLTKQLTSEG 809
Qy 320 EHAQLGIENVTDENAKKAILITDAAKDKGAAELEK--------LFKAVENLAKAAKEM-- 369
Db 810 EKLQSQIEKLRAVAAEKSVL------ESHFEELEKTLSEVKAQLKENVENAATASVKVAE 863
Qy 370 LANSVKE 376
Db 864 LTSKLQE 870

RESULT 5
US-09-708-427-19882
Sequence 19882, Application US/09708427
GENERAL INFORMATION:
    APPLICANT: N. ALEXANDROV et al.
    TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
    TITLE OF INVENTION: THEREBY
    FILE REFERENCE: 2750-1243P
    CURRENT APPLICATION NUMBER: US/09/708,427
    CURRENT FILING DATE: 2000-11-09
    NUMBER OF SEQ ID NOS: 85364
    SOFTWARE: Patentin version 3.1
SEQ ID NO 19882
    LENGTH: 1018
    TYPE: PRT
    ORGANISM: Arabidopsis thaliana
    FEATURE:
    NAME/KEY: misc_feature
    LOCATION: 1..1018
    OTHER INFORMATION:
    FEATURE:
    NAME/KEY: misc_feature
    LOCATION: 1..1018
    OTHER INFORMATION: Xaa is any amino acid
    OTHER INFORMATION: Ceres Seq. ID 1836846
US-09-708-427-19882

Query Match   9.3%; Score 172.5; DB 6; Length 1018;
Best Local Similarity   23.2%; Pred. No. 0.00018;
Matches 113; Conservative 78; Mismatches 161; Indels 135; Gaps 22;

Qy   8 KDGNTSANSADE-----------------------SVKGPNL------------TEINKKITDS-NAVLLAVKE 45
Db 405 KHGETEADSKGYLGQVAELQSTLEAFQVKSSSLEAALNIATENEKELTENLNAVTSEKKK 464

Qy   46 VEALLSSIDEIAAK------------------AIGKKIHQNNGLDTENNHNGSLLAGAYAISTL--IKQ 94
Db 465 LEA---TVDEYSVKISESENLLESIRNELNVTQG-KLESIENDLKAAGLQESEVMEKLKS 520

Qy   95 KLDGLKNEGLKEKIDAA-------------------KKCSETFTNK------LKEKH 126
Db 521 AEESLEQKG--REIDEATTKRMELEALHQSLSIDSEHRLQKAMEEFTSRDSEASSLTEKL 578

Qy 127 TDL-GKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKEL--- 182
Db 579 RDLEGKIKSYEEQLAEASGKSSSLKEKLEQTLGRLAAAESVNEKLKQEFDQAQEKSLQSS 638

Qy 183 -TSPVVAESPKK-----------PSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAV 233
Db 639 SESELLAETNQLKIKIQELEGLIGSGSVEKETALKRLEEAIE----RFNQKETESSDL 693

Qy 234 VLAVKEIETLLASIDELATKAIG-------------KKIQQNG----GLAVEA 269
Db 694 VEKLKTHENQIEEYKKLAHEASGVADTRKVELEDALSKLKNLESTIEELGAKCQGLEKES 753

Qy 270 GHNGTL------LAGAYTISKLITQKLDGLK-NSEKLKEKIENAKKCSEDFTKKL--EG 319
Db 754 GDLAEVNLKLNLELANHGSEANELQTKLSALEAEKEQTANELEASKTTIEDLTKQLTSEG 813

Qy 320 EHAQLGIENVTDENAKKAILITDAAKDKGAAELEK--------LFKAVENLAKAAKEM-- 369
Db 814 EKLQSQIEKLRAVAAEKSVL------ESHFEELEKTLSEVKAQLKENVENAATASVKVAE 867

Qy 370 LANSVKE 376
Db 868 LTSKLQE 874

RESULT 6
US-09-708-427-19881
Sequence 19881, Application US/09708427
GENERAL INFORMATION:
    APPLICANT: N. ALEXANDROV et al.
    TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
    TITLE OF INVENTION: THEREBY
    FILE REFERENCE: 2750-1243P
    CURRENT APPLICATION NUMBER: US/09/708,427
    CURRENT FILING DATE: 2000-11-09
    NUMBER OF SEQ ID NOS: 85364
    SOFTWARE: Patentin version 3.1
SEQ ID NO 19881
    LENGTH: 1269
    TYPE: PRT
    ORGANISM: Arabidopsis thaliana
    FEATURE:
    NAME/KEY: misc_feature
    LOCATION: 1..1269
    OTHER INFORMATION:
    FEATURE:
    NAME/KEY: misc_feature
    LOCATION: 1..1269
    OTHER INFORMATION: Xaa is any amino acid
    OTHER INFORMATION: Ceres Seq. ID 1836845
US-09-708-427-19881

Query Match   9.3%; Score 172.5; DB 6; Length 1269;
Best Local Similarity   23.2%; Pred. No. 0.00024;
Matches 113; Conservative 78; Mismatches 161; Indels 135; Gaps 22;

Qy   8 KDGNTSANSADE-----------------------SVKGPNL------------TEINKKITDS-NAVLLAVKE 45
Db 656 KHGETEADSKGYLGQVAELQSTLEAFQVKSSSLEAALNIATENEKELTENLNAVTSEKKK 715

Qy   46 VEALLSSIDEIAAK------------------AIGKKIHQNNGLDTENNHNGSLLAGAYAISTL--IKQ 94
Db 716 LEA---TVDEYSVKISESENLLESIRNELNVTQG-KLESIENDLKAAGLQESEVMEKLKS 771

Qy   95 KLDGLKNEGLKEKIDAA-------------------KKCSETFTNK------LKEKH 126
Db 772 AEESLEQKG--REIDEATTKRMELEALHQSLSIDSEHRLQKAMEEFTSRDSEASSLTEKL 829

Qy 127 TDL-GKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKEL--- 182
Db 830 RDLEGKIKSYEEQLAEASGKSSSLKEKLEQTLGRLAAAESVNEKLKQEFDQAQEKSLQSS 889

Qy 183 -TSPVVAESPKK-----------PSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAV 233
Db 890 SESELLAETNQLKIKIQELEGLIGSGSVEKETALKRLEEAIE----RFNQKETESSDL 944

Qy 234 VLAVKEIETLLASIDELATKAIG-------------KKIQQNG----GLAVEA 269
Db 945 VEKLKTHENQIEEYKKLAHEASGVADTRKVELEDALSKLKNLESTIEELGAKCQGLEKES 1004

Qy 270 GHNGTL------LAGAYTISKLITQKLDGLK-NSEKLKEKIENAKKCSEDFTKKL--EG 319
Db 1005 GDLAEVNLKLNLELANHGSEANELQTKLSALEAEKEQTANELEASKTTIEDLTKQLTSEG 1064

Qy 320 EHAQLGIENVTDENAKKAILITDAAKDKGAAELEK--------LFKAVENLAKAAKEM-- 369
Db 1065 EKLQSQIEKLRAVAAEKSVL------ESHFEELEKTLSEVKAQLKENVENAATASVKVAE 1118

Qy 370 LANSVKE 376
Db 1119 LTSKLQE 1125

RESULT 7

US-09-742-096-5
Sequence 5, Application US/09742096
GENERAL INFORMATION:
APPLICANT: DRUILHE, PIERRE
APPLICANT: DAUBERSIES, PIERRE
TITLE OF INVENTION: MALARIAL PRE-ERYTHROCYTIC STAGE POLYPEPTIDE MOLECULES
FILE REFERENCE: 20073USODIV
CURRENT APPLICATION NUMBER: US/09/742,096
CURRENT FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: US 08/973,642
PRIOR FILING DATE: 1998-02-06
PRIOR APPLICATION NUMBER: PCT/FR96/00894
PRIOR FILING DATE: 1996-06-12
PRIOR APPLICATION NUMBER: FR 95/07007
PRIOR FILING DATE: 1995-06-13
NUMBER OF SEQ ID NOS: 29
SOFTWARE: Patentin version 3.1

SEQ ID NO 5
LENGTH: 630
TYPE: PRT
ORGANISM: P. falciparum
US-09-742-096-5

Query Match   9.1%;   Score 168;   DB 6;   Length 630;
Best Local Similarity   18.9%;   Pred. No. 0.0002;
Matches   82; Conservative 104; Mismatches 144; Indels 104; Gaps 20;

Qy   9 DGNTS--------ANSADESVK----------------GPNLTEINKKITDSNAVLLAVKEVEALLSSID 54
            :| |:||||| |.:|:|| | :|
Db   135 DGSVASSVEESIASSVDESIDSSTEENVAPTVEEI---VAPTVEEIVAPSVVESVESVAPSVE 191

Qy   55 EIAAKAIGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKC 114
            :|:| || |: || ::| |||| | |::|||| || | :|| |:||
Db   192 ESVAPSVEESVAEN------------------VEESVAENVEEIVAPSVEESV--AENV 230

Qy   115 SETFTNKLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEEL--GKLFESV-------EVL 167
            :| | | :||| | || :| ||| | | | || |:| |||
Db   231 EESVAENVEESVAENVEESVAE-NVEESV----AENVEEIVAPTVEESVAPTVEEIV 282

Qy   168 SKAAKEMLANSVKELTSPVVAESPKKPSMVNNSGKDGNTS-ANSADESVKGPNLTE-ISK 225
            :||| | : ||: | |:|: |:| | ::||| |: :|||::|| |
Db   283 APTVEESVAPTVEEIVVPSVEES-VAPSVEESVAENVEESVAENVEESVAE-AENVEESVAE 340

Qy   226 KITESNAVVLAVKEIEFLLA-SIDELATKAIGKKIQONGGLAVEAGHNGTLLLAGAYT--I 282
            |: | :| || | | | |: ||||| | | | |||||||||| ||||||| | |
Db   283 APTVEESVAPTVEEIVVPSVEES-VAPSVEESVAENVEESVAENVEESVAE-AENVEESVAE 340

Qy   341 NVEESVA------ENVEEIVAPSVEEIVAPTVEESVAENVATNLSDNLLSNLLGGIETEEI 395
            :: | | |:| ::||||| | | | | ||
Db   341 NVEESVA------ENVEEIVAPSVEEIVAPTVEESVAENVATNLSDNLLSNLLGGIETEEI 395

Qy   283 SKLITQKLDGLKNS--------EKLK------------EKIENAKKCSEDFTKKLEGEHA 322
            :| || || || |:|: ||:|| |: | | || | |
Db   396 KDSILNEIEEVKENVVTTILEKVEETTAESVTTFSNILEEIQENTITNDTIEEKLEELHE 455

Qy   323 QL---GIENVTDENAKKAIL----------------ITDAAKDKGAAELEKLFKAVENLA 363
            |: ||| |: || | | ||||| | :|| | |:| |
Db   456 NVLSAALENTQSEEKKKEVIDVIEVKVEEVATTLIETVEQAEEESESTITEIFENLEENA 515

Qy   364 KAAKEMLANSVKEL 377
            | | | ||| |||
Db   516 VESNEKVAENLEKL 529

RESULT 8
US-09-742-096-3
Sequence 3, Application US/09742096
GENERAL INFORMATION:
APPLICANT: DRUILHE, PIERRE
APPLICANT: DAUBERSIES, PIERRE
TITLE OF INVENTION: MALARIAL PRE-ERYTHROCYTIC STAGE POLYPEPTIDE MOLECULES
FILE REFERENCE: 20073USODIV
CURRENT APPLICATION NUMBER: US/09/742,096
CURRENT FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: US 08/973,642
PRIOR FILING DATE: 1998-02-06

PRIOR APPLICATION NUMBER: PCT/FR96/00894
PRIOR FILING DATE: 1996-06-12
PRIOR APPLICATION NUMBER: FR 95/07007
PRIOR FILING DATE: 1995-06-13
NUMBER OF SEQ ID NOS: 29
SOFTWARE: Patentin version 3.1

SEQ ID NO 3
LENGTH: 1786
TYPE: PRT
ORGANISM: P. falciparum
US-09-742-096-3

Query Match   8.8%;   Score 160;   DB 6;   Length 1786;
Best Local Similarity   18.7%;   Pred. No. 0.0031;
Matches   76; Conservative 97; Mismatches 167; Indels 66; Gaps 16;

Qy   19 ESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGKKIHQNNGLDTENHNHN 78
            |: | | :||: | | | | | | ::||
Db   588 EEIVAPTVEEI---VAPTVEEIVAPSVVESVESVAPSVEESVAENVEESVAENVEESVAENVE 644

Qy   79 GSLLAG-------AYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNKLKEKHTDLGKEG 133
            |:   :| | || :|| | | |
Db   645 ESVAENVEEIVAPTVEEIVAPSVVESVESV--APSVEESVAENVEESVAENVEES 702

Qy   134 VTDADAKEAILKANGTKTKGAEELGKLFESVEVLSRAAKEMLANSVKELTSPVVAES--P 191
            | | | | | | |||||: || |: | |:||||||| | |:||
Db   703 VAE-NVEESV--AENVEESVAENV-----EEIVAPTVEEIVAPSVVESVAP 753

Qy   192 KKPSMVNNSGKDGNTSANSADESVKGPNLTE-ISKKITESNAVVLAVKEI----ETLL 244
            :| ::||| |: :|||::|| | | | | |:| | || || | :
Db   754 SVEESVEENVEE--SVAENVEESV-AENVEESVAENVEES--VAPTVEEIVAPSVEESVA 808

Qy   245 ASIDELATKAIGKKIQON-----GGLAVEAGHNGTLLLAGAYTISKLITQKLDGLKNS-- 296
            |: ||||| | | | | ||||||||| | ||||||| :|||:||| || ||
Db   809 PSVEESVAENVATNLSDNLLSNLLGGIETEETKDSIINEIEEVKENVVTTILENVEETTA 868

Qy   297 -----EKLKEKIENAKKCSEDFTKKLEGEHAQL---GIENVTDENAKKAIL---------- 339
                 |:|:|:||| |: | | || | |:| ||| |: || | |
Db   869 ESVTTFSNILEEIQENTITNDTIEEKLEELHENVLSAALENTQSEEKKEVIDVIEVKVE 928

Qy   340 ------ITDAAKDKGAAELEKLFKAVENLAKAAKEMLANSVKEL 377
                  ||||| | :|| | |:| || | | | ||| |||
Db   929 EVATTLIETVEQAEEKSANITEIFENLEENAVESNENVAENLEKL 974

RESULT 9
US-09-614-150-6213
Sequence 6213, Application US/09614150
GENERAL INFORMATION:
APPLICANT: Venter, J. Craig
APPLICANT: et al.
TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
TITLE OF INVENTION: DROSOPHILA GENES.
FILE REFERENCE: CL000728
CURRENT APPLICATION NUMBER: US/09/614,150
CURRENT FILING DATE: 2000-07-11
PRIOR APPLICATION NUMBER: 60/157,832
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: 60/160,191
PRIOR FILING DATE: 1999-10-19
PRIOR APPLICATION NUMBER: 60/161,932
PRIOR FILING DATE: 1999-10-28
PRIOR APPLICATION NUMBER: 60/164,769
PRIOR FILING DATE: 1999-11-12
PRIOR APPLICATION NUMBER: 60/173,383
PRIOR FILING DATE: 1999-12-28
PRIOR APPLICATION NUMBER: 60/175,693
PRIOR FILING DATE: 2000-01-12
PRIOR APPLICATION NUMBER: 60/184,831
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: 60/191,637

; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSEQ for Windows Version 4.0
;
; SEQ ID NO 6213
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-09-614-150-6213

Query Match   8.6%;   Score 159.5; DB 6; Length 1639;
Best Local Similarity   23.9%; Pred. No. 0.003;
Matches 105; Conservative 58; Mismatches 162; Indels 115; Gaps 19;

Qy   2 ACNNSGKDGNTSANSAD--------ESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDE   55
              :|| | :||:| | :||::|||||:|:|| | | :
Db 1107 ARDNSGDGGQTYAEVIDDLHKHLDSVR-EHLVSADKFQADANGEIDRARQNYTILDQITE 1165

Qy   56 IAAKAIGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCS 115
             ||: ||| | :||| ::|||:||||||||||||||||||:|:|::| :||||||
Db 1166 NAKKEL------QQALDLLNDEGAQALARAKEKSVEFGQQ------SEQISDISREA   1210

Qy   116 ETFTNKLK-EKHTDLGKEGVTDADAKEAILKAN--------GTKTKG--AEELGK   159
             :||:| |||: || |:| || | | :| :|||:| |||||
Db 1211 RALADKLESEAQFDLKNA----KDKADAVEKAHQLAKSAIDLQLKIGTELRSEVGLELSH 1266

Qy   160 LFESVEVLSKAAKE------------MLANSVKELTSPVVAESPKKPSMVNNSGKDGNT   206
             ||:|| :|| || |||:|||:||||||:|||:|:|||||||||
Db 1267 VKQSLGTVVQTSKEALRKANEVYDTALTLLNDVNRQTQPEIDISQLK----------KDAVA 1318

Qy   207 SANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDELATKRAIGKKIQQ---- 261
                | :|:   :: | |||:||||||||||||||:::| :|| || |
Db 1319 ANERADELLK---QITELS---NSNGELFADFETEQEL------TEALLKRAEQQQLED 1365

Qy   262 -----------NGGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSE------KL   299
                       |||||:||:||||||||:|:||||||||||:| ||
Db 1366 IELLERAKAAHDKATKRAVEQGDN-TLKEANNTYEKLAGFQSDVQRSSESAEKALQTVPNI 1424

Qy   358 AVENLAKAAKEMLANSVKEL 377
             ||: |: |:
Db 1425 EKEIQNAESLISQAEEALDGANKN----ANEAKKNAQEAQLKYAEQASKD------AELIR 1475

Qy   ....
Db 1476 RKANETKVAARNLREEAQL 1495

; RESULT 10
; US-09-708-427-3730
; Sequence 3730, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: THEREBY
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US-09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: Patentin version 3.1
;
; SEQ ID NO 3730
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..805
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..805
; OTHER INFORMATION: Ceres Seq. ID 1810877
US-09-708-427-3730

; LENGTH: 805

Query Match   8.6%;   Score 156.5; DB 6; Length 805;
Best Local Similarity   20.4%; Pred. No. 0.0019;
Matches   86; Conservative 83; Mismatches 167; Indels 85; Gaps 16;

Qy   21 VKGPNLTEINK------KITDSNAVLLAVKE----VEAL----LSSIDEIAAKAIGKKIH   66
             ||||||||| ||:| |||||||| :|| ||||||:||||:||||:
Db   50 VKHKELTEVKEAFDGLGLELENSRKKMIELEDRIRISALEAAEKLEEELQKQSASELEEKLK 109

Qy   67 QNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKI-------------DAAKKCS 115
             :||||||||||||||||||||:||:|||||||||||:|||: |||||
Db   110 IS---DERYSKTDALLSQALSQNSVLEQKLKSL--EELSRKVSELKSALIVAEEEGKKSS 164

Qy   116 --ETFTNKLKEKHTDLGKEGVTDADAKE-----AILKA-------NGTKTKGAAEELGKLFE 162
               ||:|||:|: :|| |:|||| || || |||| |:||||| |:|||:|
Db   165 IQMQEYQEKVSKLESSLNQSSARNSELEEDLRIALOKGAEHEDIGNVSTKRSVELGLFQ 224

Qy   163 SVEVLSKAAKEML------ANSVKELTSPVVAESPKKPS-------MVNNSGKDGNT 206
             ||| ||:||| ||:|||||||||:|||:||| ||||||||||
Db   225 TSQLKLEKAEEKLKDLEAIQVKNSSLEATLSVAMEKERDLSENLNAVMEKLKSSEERLEK 284

Qy   207 SANSADESVK-----------------GPNLTEISKKITESNAVVLAVKEIETLLASIDEL 250
                | |:|: |||:||||||||||:||||||||||||||:::|
Db   285 QAREIDEATTRSIELEALHKHSELKVQKTMEDFSSRDTEAKSLTERKSDLEEKIRVYEGK 344

Qy   251 ATKAIGKKIQQONGGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEK----LKEIENAKK 308
             || :||:|:|| ||||||||:|:|||||||||:| ||::||||
Db   345 LAEACGQSLSLQEELDQSSAEN-ELLADTNNQLKIKIQELEGYLDSEKETAIEKLNQKDT 403

Qy   309 CSEDFTKKLEGEHAQLGIENVTDENAKKAILITDAARDKGAAELRKLFKAVENLARAAKE 368
               |:||:| || | |||:|| |||| ||:||| | | ||| :| | |:|| |:|||
Db   404 EAKDLITKLKSH-----ENVIEEH-KRQVLEASGVADTRKVEVEEALLKLNTLESTIEE 456

Qy   369 M 369
             M
Db   457 L 457

; RESULT 11
; US-09-708-427-3729
; Sequence 3729, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
; TITLE OF INVENTION: THEREBY
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: Patentin version 3.1
;
; SEQ ID NO 3729
; TYPE: PRT
; LENGTH: 841
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..841
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..841
; OTHER INFORMATION: Ceres Seq. ID 1810876
US-09-708-427-3729

Query Match   8.6%;   Score 156.5; DB 6; Length 841;
Best Local Similarity   20.4%; Pred. No. 0.002;
Matches   86; Conservative 83; Mismatches 167; Indels 85; Gaps 16;

Qy   21 VKGPNLTEINK------KITDSNAVLLAVKE----VEAL----LSSIDEIAAKAIGKKIH   66
             ||||||||| ||:| |||||||| :|| ||||||:||||:||||:
Db   86 VKHKELTEVKEAFDGLGLELENSRKKMIELEDRIRISALEAAEKLEEELQKQSASELEEKLK 145

Qy   67 QNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKI-------------DAAKKCS 115
             :||||||||||||||||||||:||:||||||||||||:|||: |||||
Db   ...

```
Db  146 IS---DERYSKIDALLSQALSQNSVLEQKLKSL--EELSEKVSELKSALIVAEEGKKSS  200
Qy  116 ---ETFTNLKEKHTDLGKEGVTDADAKE----AILKA------NGTKTKGAEELGKLFE  162
Db  201 IQMQEYQEKVSKLESSLNQSSARNSELEEDLRIALQKGAEHEDIGNVSTKRSVELQGLFQ  260
Qy  163 SVEVLSKAAKEML----------ANSVKELTSPVVAESPKKPS------MVNNSGKDGNT  206
Db  261 TSQLKLEKAEEKLKDLEAIQVKNSSLEATLSVAMEKERDLSENLNAVMEKLKSSEERLEK  320
Qy  207 SANSADESVK----------GPNLTEISKKITESNAVVLAVKEIETLLASIDEL  250
Db  321 QAREIDEATTRSIELEALHKHSELKVQKTMEDFSSRDTEAKSLITEKSKDLEEKIRVYEGK  380
Qy  251 ATKAIGKKIQQNGGLQVAVEAGHNGTILLAGAYTISKLLITQKLIDGLKNSEK--LKEKIENAKK  308
Db  381 LAEACGQSLSLQEELDQSSAEN-ELLADTNNQLKIKIQELEGYLDSEKETAIEKLNQKDT  439
Qy  309 CSEDFTKKLEGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLAKAAKE  368
Db  440 EAKDLITKLKSH-----ENVIEEH-KRQVLEASGVADTRKVEVEEALLKNTLESTIEE  492
Qy  369 M  369
Db  493 L  493
```

RESULT 12
US-09-708-427-32624
Sequence 32624, Application US/09708427
GENERAL INFORMATION:
APPLICANT: N. ALEXANDROV et al.
TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
TITLE OF INVENTION: THEREBY
FILE REFERENCE: 2750-1243P
CURRENT APPLICATION NUMBER: US/09/708,427
CURRENT FILING DATE: 2000-11-09
NUMBER OF SEQ ID NOS: 85364
SOFTWARE: Patentin version 3.1
SEQ ID NO 32624
LENGTH: 1038
TYPE: PRT
ORGANISM: Arabidopsis thaliana
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1..1038
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: 1..1038
OTHER INFORMATION: Ceres Seq. ID 1834662
US-09-708-427-32624

Query Match   8.4%;   Score 155.5;   DB 6;   Length 1038;
Best Local Similarity   25.0%;   Pred. No. 0.0032;
Matches   104;   Conservative   53;   Mismatches   186;   Indels   73;   Gaps   16;

```
Qy    7 GKDGNTSAN---SADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGK   63
Db  372 GKEGNFECNVLSDAEEKSSVINIPVANHLQEPRN---IPVKLSENHLPKPTEPT-----K  423
Qy   64 KIHQNNGLDTENNHNGSLLAGA------YAISTLIKQKLDGLKNEGLKEKIDAAK--KC  114
Db  424 RIAKNEPVKSTKKEQSSSSSEASKLPPVCLRVDPLPKERNGGSKSVSHPKRMEKSKETKI  483
Qy  115 SETFTNKLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEM  174
Db  484 AAPLSSKKAESRTVPEACNVKCEDANAEMKMAEGSLNALRTEKG----SVESNSLQEES  539
Qy  175 LANSVKELTSPVVAESPKKPSMVNNSGKDGNTSANSADESVKGPNLT--EISKKITESNA  232
Db  540 NGEIIKPCEAKENREQPAKSFTEEE------AARIIQSMYRGYDVRRWEPIKKLKEIAT  593
Qy  233 VVLAV----KEIETLLASIDELATKAIGKKIQQNGGL------AVEAGHNG--TLLA  277
Db  594 VREQMGDVVKRIEALEASTDQHIEE--KEIVVNGELVMNLLLKLDAVEGLHPSIREFRK  650
Qy  278 GAYTISKLLITQKLIDGLKNS-----EKLKEKIENAKKCCSEDFTKKLEGEHAQLGIEN--V  329
Db  651 ALATELSSIQDKLDSLKNSCASAEKEAVKEQVELKSQPSD---SPVNLEHSQLTEENKMV  707
Qy  330 TDENAKKAILI-------------TDAAKDKGAAELEKLFKAVENLAKAAKEMLANS  373
Db  708 SDTNLEKVLRLSPEEHPMSVLNRFDEKQAESAAETEEGYGLFETLATDSKQATENA  763
```

RESULT 13
US-09-708-427-32623
Sequence 32623, Application US/09708427
GENERAL INFORMATION:
APPLICANT: N. ALEXANDROV et al.
TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
TITLE OF INVENTION: THEREBY
FILE REFERENCE: 2750-1243P
CURRENT APPLICATION NUMBER: US/09/708,427
CURRENT FILING DATE: 2000-11-09
NUMBER OF SEQ ID NOS: 85364
SOFTWARE: Patentin version 3.1
SEQ ID NO 32623
LENGTH: 1042
TYPE: PRT
ORGANISM: Arabidopsis thaliana
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1..1042
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: 1..1042
OTHER INFORMATION: Ceres Seq. ID 1834661
US-09-708-427-32623

Query Match   8.4%;   Score 155.5;   DB 6;   Length 1042;
Best Local Similarity   25.0%;   Pred. No. 0.0032;
Matches   104;   Conservative   53;   Mismatches   186;   Indels   73;   Gaps   16;

```
Qy    7 GKDGNTSAN---SADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGK   63
Db  376 GKEGNFECNVLSDAEEKSSVINIPVANHLQEPRN---IPVKLSENHLPKPTEPT-----K  427
Qy   64 KIHQNNGLDTENNHNGSLLAGA------YAISTLIKQKLDGLKNEGLKEKIDAAK--KC  114
Db  428 RIAKNEPVKSTKKEQSSSSSEASKLPPVCLRVDPLPKERNGGSKSVSHPKRMEKSKETKI  487
Qy  115 SETFTNKLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEM  174
Db  488 AAPLSSKKAESRTVPEACNVKCEDANAEMKMAEGSLNALRTEKG----SVESNSLQEES  543
Qy  175 LANSVKELTSPVVAESPKKPSMVNNSGKDGNTSANSADESVKGPNLT--EISKKITESNA  232
Db  544 NGEIIKPCEAKENREQPAKSFTEEE------AARIIQSMYRGYDVRRWEPIKKLKEIAT  597
Qy  233 VVLAV----KEIETLLASIDELATKAIGKKIQQNGGL------AVEAGHNG--TLLA  277
Db  598 VREQMGDVVKRIEALEASTDQHIEE--KEIVVNGELVMNLLLKLDAVEGLHPSIREFRK  654
Qy  278 GAYTISKLLITQKLIDGLKNS-----EKLKEKIENAKKCCSEDFTKKLEGEHAQLGIEN--V  329
Db  655 ALATELSSIQDKLDSLKNSCASAEKEAVKEQVEIKSQPSD---SPVNLEHSQLTEENKMV  711
Qy  330 TDENAKKAILI-------------TDAAKDKGAAELEKLFKAVENLAKAAKEMLANS  373
Db  712 SDINLEKVLRLSPEEHPMSVLNRFDEKQAESAAETEECYGLFETLATDSKQATENA  767
```

RESULT 14

US-09-708-427-32622
; Sequence 32622, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: Patentin version 3.1
; SEQ ID NO: 32622
; LENGTH: 1043
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..1043
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..1043
; OTHER INFORMATION: Ceres Seq. ID 1834660
US-09-708-427-32622

Query Match   8.4%; Score 155.5; DB 6; Length 1043;
Best Local Similarity   25.0%; Pred. No. 0.0032;
Matches 104; Conservative 53; Mismatches 186; Indels 73; Gaps 16;

```
Qy    7 GKDGNTSAN---SADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGK  63
          |:|:|       |:|       |:|  ||||   :  :: :|| ::| :|  : |:|
Db  377 GKEGNFECNVLSDAEERSSVINIPVANHLQEPRN---IPVKLSENHLPKPTEPT----K 428

Qy   64 KIHQNGLDTENNHNGSLLAGA------YAISTLIKQKLDGLKNEGLIKEKIDAAK--KC 114
Db  429 RIAKNEPVKSTKKEQSSSSSEASKLPPVCLRVDPLPKERNGGSKSVSHPKRMEKSKETKI 488

Qy  115 SETFTNKLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEM 174
Db  489 AAPLSSKKAESRTVPEACNVKCEDANAEMKMAEGSLNALRTEKG----SVESNSNLQEES 544

Qy  175 LANSVKELTSPVVAESPKKPSMVNNSGKDGNTSANSADESVKGPNLT---EISKKITESNA 232
Db  545 NGEIIKPCEAKENREQPAKKSFTEEE------AARIQSMYRGYDVRRWEPIKPLKEIAT 598

Qy  233 VVLAV---KEIETLLASIDELATKAIGKKIQQNGGL--------AVEAGHNG--TLLA 277
Db  599 VREQMGDVKKRTEALRASTDQHEE---KEIVVNGELVMNLLKLDAVEGLHPSIREFRK 655

Qy  278 GAYTISKLITQKLDGLKNS----EKLKEKIENAKKCCSEDFTKKLEGEHAQLGIEN--V 329
Db  656 ALATELSSIQDKLDSLKNSCASAAEKEAVKEQVEIKSQPSD---SPVNLEHSQLTEENKMV 712

Qy  330 TDENAKKAILI----------TDAAKDKGAAELEKLFKAVENLAKAAKEMLANS 373
Db  713 SDTNLEKVLRLSPEEHPMSVLNRTDEKQAESAAETEECYGLFETLATDSKQATENA 768
```

RESULT 15
US-10-032-585-7646
; Sequence 7646, Application US/10032585
; GENERAL INFORMATION:
; APPLICANT: Terry, Roemer D.
; APPLICANT: Bo, Jiang
; APPLICANT: Charles, Boone
; APPLICANT: Howard, Bussey
; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
; FILE REFERENCE: 10182-005-999
; CURRENT APPLICATION NUMBER: US/10/032,585
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 8000
; SOFTWARE: Patentin version 3.1
; SEQ ID NO: 7646

; LENGTH: 1881
; TYPE: PRT
; ORGANISM: Candida albicans
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (1881)..(1881)
; OTHER INFORMATION: X=any amino acid
US-10-032-585-7646

Query Match   8.3%; Score 153; DB 7; Length 1881;
Best Local Similarity   23.8%; Pred. No. 0.011;
Matches 91; Conservative 59; Mismatches 146; Indels 86; Gaps 16;

```
Qy   27 TEINKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGKKIH--QNNGLDTENNHNGSLLAG  84
Db 1529 TELKDRISE------VEKERAMLSENSETVIKEYSDKIRSLESKINSIKENHSKE---- 1577

Qy   85 AYAISTLIKQKLDGLKNE--GLKEKIDAAKKCSETFTNLKEKHTDLGKEGVTDADAKEA 142
Db 1578 ---ITTHNEQK-TSLKQDIAKLSQDHESAQTQLEDKENQLKELKASLEKHNTESATSIE- 1632

Qy  143 ILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKELTSPVVAESPKKPSMVNNSGK 202
Db 1633 ------EKNNQIKELSETIKSLKTELKTSGDALKQSQKEY----------KTLKIKNSDT 1676

Qy  203 DG-------NTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETL----LAS 246
Db 1677 ESKLEKQLEELEKVKSDLQTADEKLKGITEREIALK-------------SELETVKNSGLST 1725

Qy  247 IDELA--TKAIGKKIQQNGGLAVERGHNGTLLAGAYTISKLITQKLDGLKN--SEKLKEK 302
Db 1726 TSELAALTKTPVKSLEKEKEELQFLSGNKSKELEDYIQKHSDISEKLKALTDELKEKTKQF 1785

Qy  303 IENAKKCSE---DFT---KKLEGEHAQLG-IENVTDENAKKAILITDAAKDKGAAELEKL 355
Db 1786 DDSKKKLITELENDLITSTKKELETEKTQTSKFKNLEER-------------KDKEIVKLNKE 1833

Qy  356 FKAVENLAKAAKEMLANSVKEL 377
Db 1834 LELLKNDNSGAKKELSEKVSKL 1855
```

Search completed: March 18, 2002, 09:58:21
Job time: 379 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 09:56:47 ; Search time 68.77 Seconds
(without alignments)
417.592 Million cell updates/sec

Title: US-09-596-746A-30
Perfect score: 1850
Sequence: 1 MACNNSGKDGNTSANSADES..........AVENIAKAAKEMLANSVKEL 377

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_68:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	945	51.1	210	2	G70218	outer surface prot
2	938	50.7	210	2	S69927	outer surface prot
3	876	47.4	193	2	S70287	outer surface prot
4	863	46.6	191	2	S69916	outer surface prot
5	853	46.1	191	2	I40153	outer surface prot
6	826	44.6	177	2	I40129	outer surface prot
7	823	44.5	178	2	I40145	outer surface prot
8	774.5	41.9	178	2	I40104	outer surface prot
9	773.5	41.8	211	2	S70279	outer surface prot
10	739	39.9	193	2	I40145	outer surface prot
11	709	38.3	178	2	I40124	outer surface prot
12	695.5	37.6	210	2	I40144	outer surface prot
13	692.5	37.4	211	2	S69918	outer surface prot
14	690.5	37.3	193	2	I40277	outer surface prot
15	686.5	37.1	211	2	S70280	outer surface prot
16	686	37.1	190	2	I40068	outer surface prot
17	685.5	37.1	209	2	S70273	outer surface prot
18	683.5	36.9	191	2	I40273	outer surface prot
19	675	36.5	194	2	S70278	outer surface prot
20	674.5	36.5	209	2	S70277	outer surface prot
21	670	36.2	212	2	S69917	outer surface prot
22	666	36.0	193	2	S70254	outer surface prot
23	664.5	35.9	193	2	S70276	outer surface prot
24	664.5	35.9	211	2	S70274	outer surface prot
25	664	35.9	212	2	S69922	outer surface prot
26	662.5	35.8	191	2	S70284	outer surface prot
27	662.5	35.8	193	2	S70265	outer surface prot
28	662	35.8	194	2	S70268	outer surface prot
29	661.5	35.8	211	2	S69932	outer surface prot
30	657.5	35.5	209	2	I40142	outer surface prot
31	655	35.4	212	2	I40279	outer surface prot
32	650	35.1	212	2	I40143	outer surface prot
33	646	34.9	214	2	S69916	outer surface prot
34	641.5	34.7	193	2	S70286	outer surface prot
35	638	34.5	194	2	S70289	outer surface prot
36	635.5	34.4	191	2	S70288	outer surface prot
37	635.5	34.4	203	2	I40108	outer surface prot
38	631	34.1	212	2	S20543	outer surface prot
39	630.5	34.1	209	2	I40281	outer surface prot
40	630.5	34.1	212	2	I40285	outer surface prot
41	630	34.1	192	2	S70285	outer surface prot
42	628	33.9	210	2	S69925	outer surface prot
43	627	33.9	210	2	S69923	outer surface prot
44	627	33.9	212	2	S69921	outer surface prot
45	622.5	33.6	209	2	S69926	outer surface prot

ALIGNMENTS

RESULT 1
G70218

outer surface protein C - Lyme disease spirochete
(Borrelia burgdorferi (Lyme disease spirochete))

C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 21-Jul-2000
C;Accession: G70218; I40269; S37726; S70281
A;Molecule type: DNA
A;Residues: 1-210 <RES>
A;Cross-references: GB:D49497; NID:g707092; PIDN:BAA08457.1; PID:g769684
R;Jauris-Heipke, S.; Fuchs, R.; Motz, M.; Preac-Mursic, V.; Schwab, E.; Soutschek, E.
; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997
A;Reference number: A70100; MUID:98065943
A;Accession: G70218
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-210 <KLE>
A;Cross-references: GB:AE000792; NID:g3253098; PIDN:AAC66329.1; PID:g2689901; TGR:BB
A;Experimental source: strain B31
A;Accession: G70218
R;Fukunaga, M.; Hamase, A.
J. Clin. Microbiol. 33, 2415-2420, 1995
A;Title: Outer surface protein C gene sequence analysis of Borrelia burgdorferi sensu
A;Reference number: I40269; MUID:96025162
A;Accession: I40269
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-210
A;Cross-references: EMBL:X69596; NID:g311391; PIDN:CAA49306.1; PID:g311392
R;Livey, I.; Gibbs, C.P.; Schuster, R.; Dorner, F.
Mol. Microbiol 18, 257-269, 1995
A;Title: Evidence for lateral transfer and recombination in OspC variation in Lyme di
A;Reference number: S70255; MUID:96296448
A;Accession: S70281
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 19-210 <LIV>
A;Cross-references: EMBL:L42887; NID:g858715; PIDN:AAB36995.1; PID:g1695212
C;Genetics:
A;Gene: ospC
C;Superfamily: Lyme disease spirochete surface protein C

RESULT 2
S69927
outer surface protein C precursor - Lyme disease spirochete (strain Pka)
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Variety: strain Pka
A;Date: 15-Feb-1997 #sequence_revision 27-Feb-1997 #text_change 26-May-2000
C;Accession: S69927; S72669
R;Jauris-Helpke, S.; Liegl, G.; Preac-Mursic, V.; Roessler, D.; Schwab, E.; Soutschek, E.
J. Clin. Microbiol. 33, 1860-1866, 1995
A;Title: Molecular analysis of genes encoding outer surface protein C (OspC) of Borrella
A;Accession: S69927
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-193 <LIV>
A;Cross-references: EMBL:X69589
A;Experimental source: strain Pka
A;Residues: 1-210 <JAU>
A;Cross-references: EMBL:X69589
A;Experimental source: strain Pka
R;Jauris, S.
submitted to the EMBL Data Library, February 1994
A;Reference number: S72669
A;Accession: S72669
A;Molecule type: DNA
A;Residues: 1-124,'D',126-139,'E',141-210 <JAW>
A;Cross-references: EMBL:X69589
Gene: ospC
C;Genetics:
Superfamily: Lyme disease spirochete surface protein C

Query Match   51.1%; Score 945; DB 2; Length 210;
Best Local Similarity 97.8%; Pred. No. 5, 5e-39;
Matches 190; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy   1 MACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKA 60
           ::|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   17 ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA 76

Qy   61 IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSEFFTN 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   77 IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSEFFTN 136

Qy   121 KLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   137 KLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 196

Qy   181 ELTSPVVAESPKKP 194
           ||||||||||||||
Db   197 ELTSPVVAESPKKP 210

RESULT 3
S70287
outer surface protein C - Lyme disease spirochete
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 26-May-2000
C;Accession: S70287
R;Livey, I.; Gibbs, C.P.; Schuster, R.; Dorner, F.
Mol. Microbiol. 18, 257-269, 1995
A;Title: Evidence for lateral transfer and recombination in OspC variation in Lyme di
A;Reference number: S70255; MUID:96296448
A;Accession: S70287
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-193 <LIV>
A;Cross-references: EMBL:L42895; NID:g858723; PIDN:AAB37003.1; PID:g1695220
A;Experimental source: strain 28354
A;Gene: ospC
C;Superfamily: Lyme disease spirochete surface protein C

Query Match   47.4%; Score 876; DB 2; Length 193;
Best Local Similarity 100.0%; Pred. No. 1e-35;
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   198 NNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDELATKAIGKK 257
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2 NNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDELATKAIGKK 61

Qy   258 KIQQNGGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKEKIENAKKCSEDFTKKL 317
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   62 KIQQNGGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKEKIENAKKCSEDFTKKL 121

Qy   318 EGEHAQLGIENVTDENAKKAILITDAAKDGAAELEKLFKAVENLAKAAKEMLANSVKEL 377
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   122 EGEHAQLGIENVTDENAKKAILITDAAKDGAAELEKLFKAVENLAKAAKEMLANSVKEL 181

RESULT 4
I40153
outer surface protein C - Lyme disease spirochete (fragment)
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 26-May-2000
C;Accession: I40153
R;Fung, B.P.; McHugh, G.L.; Leong, J.M.; Steere, A.C.
Infect. Immun. 62, 3213-3221, 1994
A;Title: Humoral immune response to outer surface protein C of Borrelia burgdorferi i
A;Reference number: I40153; MUID:94314437
A;Accession: I40153
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-191 <RES>
A;Cross-references: EMBL:U08284; NID:g469561; PIDN:AAA21460.1; PID:g469562
A;Gene: ospC
C;Genetics:
A;Superfamily: Lyme disease spirochete surface protein C

Query Match   46.6%; Score 863; DB 2; Length 191;
Best Local Similarity 99.4%; Pred. No. 4.2e-35;
Matches 178; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy   199 NSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDELATKAIGKK 258
           ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 NSGKGGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDELATKAIGKK 60

Qy   259 IQQNGGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKEKIENAKKCSEDFTKKLE 318
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 IQQNGGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKEKIENAKKCSEDFTKKLE 120

Qy   319 GEHAQLGIENVTDENAKKAILITDAAKDGAAELEKLFKAVENLAKAAKEMLANSVKEL 377
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GEHAQLGIENVTDENAKKAILITDAAKDGAAELEKLFKAVENLAKAAKEMLANSVKEL 179

RESULT 5
I40129
outer surface protein C - Lyme disease spirochete (strain TXGW) (fragment)
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 26-May-2000
C;Accession: I40129; S54199
R;Thelsen, M.; Borre, M.; Mathiesen, M.J.; Mikkelsen, B.; Lebech, A.M.; Hansen, K.
J. Bacteriol. 177, 3036-3044, 1995
A;Title: Evolution of the Borrelia burgdorferi outer surface protein OspC.
A;Reference number: I40129
A;Accession: I40129
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-177 <RES>
A;Cross-references: EMBL:X84783; NID:g793825; PIDN:CAA59254.1; PID:g793826
A;Genetics:
A;Gene: OspC
C;Superfamily: Lyme disease spirochete surface protein C

Query Match   46.1%; Score 853; DB 2; Length 177;
Best Local Similarity 98.3%; Pred. No. 1.1e-34;
Matches 174; Conservative   1; Mismatches   2; Indels   0; Gaps   0;

Qy   11 NTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGKKIHQNNG 70
         ||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db   1 NTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGKKIHQNNG 60

Qy   71 LDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNKLKEKHTDLG 130
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 LDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNKLKEKHTDLG 120

Qy 131 KEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKELTSPVV 187
         | ||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db 121 KRGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKELTSPVV 177

RESULT 6
I40125
outer surface protein C - Lyme disease spirochete (strain MUL) (fragment)
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 26-May-2000
C;Accession: I40125; S54195
R;Thelsen, M.; Borre, M.; Mathiesen, M.J.; Mikkelsen, B.; Lebech, A.M.; Hansen, K.
J. Bacteriol. 177, 3036-3044, 1995
A;Title: Evolution of the Borrelia burgdorferi outer surface protein OspC.
A;Reference number: I40125
A;Accession: I40125
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-178 <RES>
A;Cross-references: EMBL:X84779; NID:g793817; PIDN:CAA59250.1; PID:g793818
A;Genetics:
A;Gene: OspC
C;Superfamily: Lyme disease spirochete surface protein C

Query Match   44.6%; Score 826; DB 2; Length 178;
Best Local Similarity 98.8%; Pred. No. 2.3e-33;
Matches 171; Conservative   0; Mismatches   2; Indels   0; Gaps   0;

Qy 205 NTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDELATRAIGKKIQQNGG 264
         |||||||||||||||||||||||| |||| |||||| |||||||||||| |||||| |||
Db   1 NTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDELATRAIGKKIQQNGG 60

Qy 265 LAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKEKIENAKKCSEDFTKKLEGEHAQL 324
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 LAVEAGHNGTLSLAGAYTISKLITQKLDGLKNSEKLKEKIENAKKCSEDFTKKLEGEHAQL 120

Qy 325 GIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLAKAAKEMLANSVKEL 377
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 GIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLPKAAKEMLANSVKEL 173

RESULT 7

RESULT 8
I40145
outer surface protein C - Lyme disease spirochete (strain 272) (fragment)
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 26-May-2000
C;Accession: I40145; S54184
R;Thelsen, M.; Borre, M.; Mathiesen, M.J.; Mikkelsen, B.; Lebech, A.M.; Hansen, K.
J. Bacteriol. 177, 3036-3044, 1995
A;Title: Evolution of the Borrelia burgdorferi outer surface protein OspC.
A;Reference number: I40145
A;Accession: I40145
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-178 <RES>
A;Cross-references: EMBL:X84785; NID:g793787; PIDN:CAA59256.1; PID:g793788
A;Genetics:
A;Gene: OspC
C;Superfamily: Lyme disease spirochete surface protein C

Query Match   44.5%; Score 823; DB 2; Length 178;
Best Local Similarity 98.3%; Pred. No. 3.0e-33;
Matches 170; Conservative   1; Mismatches   2; Indels   0; Gaps   0;

Qy 205 NTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDELATKAIGKKIQQNGG 264
         |||||||||||||||||||||||| |||| ||||||||||||||||||| |||||| |||
Db   1 NTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGKKIQQNGG 60

Qy 265 LAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKEKIENAKRCSEDFTKKLEGEHAQL 324
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 LAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKEKIENAKRCSEDFTKKLEGEHAQL 120

Qy 325 GIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLAKAAKEMLANSVKEL 377
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 GIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLPKAAKEMLANSVKEL 173

RESULT 8
I40104
outer surface protein C precursor - Lyme disease spirochete
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 21-Jul-2000
C;Accession: I40145
R;Stevenson, B.; Barthold, S.W.
FEMS Microbiol. Lett. 124, 367-372, 1994
A;Title: Expression and sequence of outer surface protein C among North American isol
A;Reference number: I40143; MUID:95154673
A;Accession: I40145
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-211 <RES>
A;Cross-references: EMBL:U04282; NID:g2314881; PIDN:AAC45540.1; PID:g434666
A;Genetics:
A;Gene: OspC
C;Superfamily: Lyme disease spirochete surface protein C

Query Match   41.9%; Score 774.5; DB 2; Length 211;
Best Local Similarity 81.0%; Pred. No. 8.1e-31;
Matches 158; Conservative   14; Mismatches   22; Indels   1; Gaps   1;

Qy   1 MACNNSGKDGN-TSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAK 59
         :||:| |||| |||||||||||||||||||||||||||||||||||||||||||||:|
Db   17 ISCNNSGKDGNAASTNPADESVKGPNLTEISKKITDSNAVLLAVKEVGALLTSIDELATK 76

Qy   60 AIGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFT 119
         |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db   77 AIGKKIHQNNGLDTENNHNGSLLAGAYAISTLITQKLDGLKNEELKEKIAAVKKCSEEFT 136

Qy 120 NKLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSV 179
         |||| ||||||:|||||||||||||::||:||||||||||||||||||||||||||||||
Db 137 NKLKSSHTELGKQDAQDDDAKKAILRTHNTKDKGAEELDKLFKAVENLSKAAKEMLSNSV 196

Qy 180 KELTSPVVAESPKKP 194

```
              |||||||||||||||
Db  197 KELTSPVVAESPKKP 211
```

RESULT 9
S70279
outer surface protein C - Lyme disease spirochete
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 26-May-2000
C;Accession: S70279
R;Livey, I.; Gibbs, C.P.; Schuster, R.; Dorner, F.
Mol. Microbiol. 18, 257-269, 1995
A;Title: Evidence for lateral transfer and recombination in OspC variation in Lyme disea
A;Reference number: S70255; MUID:96296448
A;Accession: S70279
C;Genetics:
A;Gene: ospc
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-193 <LIV>
cross-references: EMBL:L42898; NID:g858729; PIDN:AAB37007.1; PID:g1695223
A;Experimental source: strain 25015
C;Genetics:
A;Status: preliminary
C;Superfamily: Lyme disease spirochete surface protein C

Query Match   41.8%; Score 773.5; DB 2; Length 193;
Best Local Similarity 81.9%; Pred. No. 8.2e-31;
Matches 158; Conservative 11; Mismatches 23; Indels 1; Gaps 1;

```
Qy   3 CNNSGKDGN-TSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAKAI 61
         ||||||||| :||||||||||||||||| ||||||||::||||||||:||||:||||
Db   1 CNNSGKDGNAASTNPADESVKGPNLTEISKKITDSNTVLAVKEVEALLTSIDELATKAI 60

Qy  62 GKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNK 121
         |||||||||||||||||||||||||| ||||  ||| || ||||:|| ||||| ||||
Db  61 GKKIHQNNGLDTENNHNGSLLAGAYAISTLITQKLGGLKNEELKEKIAAVKKCSEEFTNK 120

Qy 122 LKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 181
         ||  || ||  |  ||| |||| :: |  | ||||| ||| ||| ||||||||||:||:
Db 121 LKSSHTELGKQDAQDDDAKKAILRTHNTKDKGAEELDKLFKPVENLSKAAKEMLSNSVKE 180

Qy 182 LTSPVVAESPKKP 194
         |||||||||||||
Db 181 LTSPVVAESPKKP 193
```

● ULT 10
124
outer surface protein C - Lyme disease spirochete (strain KIPP) (fragment)
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 26-May-2000
C;Accession: I40124; S54194
R;Theisen, M.; Borre, M.; Mathiesen, M.J.; Mikkelsen, B.; Lebech, A.M.; Hansen, K.
J. Bacteriol. 177, 3036-3044, 1995
A;Title: Evolution of the Borrelia burgdorferi outer surface protein OspC.
A;Reference number: I40124; MUID:95286481
A;Accession: I40124
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Residues: 1-178 <RES>
A;Cross-references: EMBL:X84782; NID:g793815; PIDN:CAA59253.1; PID:g793816
A;Molecule type: DNA
C;Superfamily: Lyme disease spirochete surface protein C

Query Match   39.9%; Score 739; DB 2; Length 178;
Best Local Similarity 89.0%; Pred. No. 3.3e-29;
Matches 154; Conservative 6; Mismatches 13; Indels 0; Gaps 0;

```
Qy 205 NTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDELATRAIGKKIQONGG 264
         |||||||||||||||||||||||||||||   |||||||||||||||||||||||||||
Db   1 NTSANSADESVKGPNLTEISKKITESNAGGLAVKEIETLLASIDELATRAIGKKIQONGG 60

Qy 265 LAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKEKIENAKKCSEDFTKKLEGEHAQL 324
```

```
              |||||||||||||||  ||||||||||||||||||| : |||||: ||||||||||||
Db  61 LAVEAGHNGTSLAGAYTISKLITQKIRWIENSEKITGKIENAKKCSEDFTNKLEGEHAQL 120
```

```
Qy 325 GIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLAKAAAKEMLANSVKEL 377
         ||||:||||| ||||:||||||||||| |||||||||||   ||||||||||||
Db 121 GTENVTDENADKAILTTDAAKDKGAAELAKLFKAVENLPQAAKEMLANSVKEL 173
```

RESULT 11
I40144
outer surface protein C precursor - Lyme disease spirochete
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 21-Jul-2000
C;Accession: I40144; S70282
R;Livey, I.; Gibbs, C.P.; Schuster, R.; Dorner, F.
Mol. Microbiol. 18, 257-269, 1995
A;Title: Evidence for lateral transfer and recombination in OspC variation in Lyme d1
A;Reference number: S70255; MUID:96296448
A;Accession: I40144; S70282
A;Molecule type: DNA
A;Residues: 19-210 <LIV>
A;Cross-references: EMBL:L42893; NID:g858721; PIDN:AAB37001.1; PID:g1695218
A;Experimental source: strain 297
C;Genetics:
A;Gene: ospC
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Residues: 1-210 <RES>
A;Cross-references: EMBL:U04281; NID:g434663; PIDN:AAC43297.1; PID:g434664
R;Livey, I.; Gibbs, C.P.; Schuster, R.; Dorner, F.
Mol. Microbiol. 18, 257-269, 1995
A;Title: Expression and sequence of outer surface protein C among North American isol
A;Reference number: I40143; MUID:95154673
A;Accession: I40144
A;Status: preliminary
A;Molecule type: DNA
C;Superfamily: Lyme disease spirochete surface protein C

Query Match   38.3%; Score 709; DB 2; Length 210;
Best Local Similarity 75.4%; Pred. No. 1.1e-27;
Matches 147; Conservative 22; Mismatches 24; Indels 2; Gaps 2;

```
Qy   1 MACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAKKA 60
         : ||||||||||||||||||||||||| ||||:||| |||:|||||||  |:|||| |||
Db  17 ISCNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVETTLTSIDEL-AKA 75

Qy  61 IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEG-LKEKIDAAKKCSETFT 119
         :||||: ::  |:  ||||| ||||: |||||  |:  |:   |||||:||||||| ||
Db  76 IGKKIKNDVSLDNEADHNGSLISGAYLISTLITKISAIKDSGELKAEIEKAKKCSEEFT 135

Qy 120 NKLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSV 179
         ||| ||||||||||||||: |:|||| :::|:|||| ||| ||||| :|||||||| :||
Db 136 AKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLSKAAKEMLTNSV 195

Qy 180 KELTSPVVAESPKKP 194
         |||||||||||||||
Db 196 KELTSPVVAESPKKP 210
```

RESULT 12
S69918
outer surface protein C precursor - Lyme disease spirochete (strain PBre)
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
A;Variety: strain PBre
C;Date: 06-Dec-1996 #sequence_revision 14-Feb-1997 #text_change 26-May-2000
C;Accession: S69918; S72674; I40103
R;Jauris-Heipke, S.; Liegl, G.; Preac-Mursic, V.; Roessler, D.; Schwab, E.; Soutschek
J. Clin. Microbiol. 33, 1860-1866, 1995
A;Title: Molecular analysis of genes encoding outer surface protein C (OspC) of Borre
A;Reference number: I40047; MUID:95395018
A;Accession: S69918
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA

A;Residues: 1-211 <JAU>
A;Cross-references: EMBL:X81522; NID:g872021; PIDN:CAA57242.1; PID:g872022
R;Roesler, D.
A;Experimental source: strain PBre
A;Reference number to the EMBL Data Library, September 1994
submitted
A;Accession: S72674
A;Molecule type: DNA
A;Residues: 1-152, 'E',154-211 <ROE>
A;Cross-references: EMBL:X81522; NID:g872021; PIDN:CAA57242.1; PID:g872022
A;Genetics:
A;Gene: ospC
A;Superfamily: Lyme disease spirochete surface protein C

Query Match   37.6%; Score 695.5; DB 2; Length 211;
Best Local Similarity   77.6%; Pred. No. 4.9e-27;
Matches 152; Conservative 12; Mismatches 29; Indels 3; Gaps 3;

Qy   1 MACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKA 60
            ::||||||| |||||||||||||||||||||||||||||||||||||||||||||:||| |||
Db   17 ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEI-AKA 75

Qy   61 IGKKIHQNNGLDTENNNGNGSLLAGAYAISTLIKQKLDGLK-NEGLKEKIDAARKCSETFT 119
            |||| ||||||||:|||||||||||||||||:||||||||||| :||||||||||||:|: |||
Db   76 IGKKINDGSLGDEANHNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKCSEEFS 135

Qy   120 NKLKEKHTDLGKEGVTDADAKEAILKANGT-KTKGAEELGKLFESVEVLSKAAKEMLANS 178
            ||| || ||||||| ||| |||||||||||| ||:||||||||||||||||:||||||||:|
Db   136 TKLKDNHAQLGIGQVTDDNAKKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANS 195

Qy   179 VKELTSPVVAESPKKP 194
            ||||||||||||||||
Db   196 VKELTSPVVAESPKKP 211

●

RESULT 13
I40277
outer surface protein C precursor - Borrelia afzelii
C;Species: Borrelia afzelii
A;Accession: I40277
C;Date: 04-Sep-1997 #sequence_revision 04-Sep-1997 #text_change 26-May-2000
A;Molecule type: DNA
R;Fukunaga, M.; Hamase, A.
J. Clin. Microbiol. 33, 2415-2420, 1995
A;Title: Outer surface protein C gene sequence analysis of Borrelia burgdorferi sensu la
A;Reference number: I40277
A;Accession: I40277
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA <RES>
Residues: 1-211 <RES>
Cross-references: GB:D49501; NID:g707096; PIDN:BAA08461.1; PID:g769688
Superfamily: Lyme disease spirochete surface protein C

Query Match   37.4%; Score 692.5; DB 2; Length 211;
Best Local Similarity   74.5%; Pred. No. 6.9e-27;
Matches 146; Conservative 20; Mismatches 27; Indels 3; Gaps 3;

Qy   1 MACNNSGKDGNT-SANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAK 59
            ::||||||| ||| |||||||||||||||||||||||||||||||||||||||||:||||
Db   17 ISCNNSGKGGDSASTNPTDESAKGPNLTEISKKITDSNAFVLAVKEVETLVSSIDELATK 76

Qy   60 AIGKKIHQNNGLDTENNNGSLLAGAYAISTLIKQKLDGLKN-EGLKEKIDAARKCSETFT 118
            ||||||||||||| |||| |:|||||||||||||||:||||| :||||||||:||||:
Db   77 AIGQKI-ONNGLAAEADKNGSLLAGAYAISTLIKQKLDGLKGLEGLNKEIAEARKCSEAF 135

Qy   119 TNKLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANS 178
            |||||| ||||| |||||| | | || | |||:||||||||||:||:||:||||||||||
Db   136 TKKLQDSNADLGRHNATDADSKEAILKNGTKRGAKELEELFKSVESLSKAAKEALSNS 195

Qy   179 VKELTSPVVAESPKKP 194
            ||||||||||||||||
Db   196 VKELTSPVVAESPKKP 211

RESULT 14
S70280
outer surface protein C - Lyme disease spirochete
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 26-May-2000
C;Accession: S70280
R;Livey, I.; Gibbs, C.P.; Schuster, R.; Dorner, F.
Mol. Microbiol. 18, 257-269, 1995
A;Title: Evidence for lateral transfer and recombination in OspC variation in Lyme di
A;Reference number: S70255; MUID:96296448
A;Accession: S70280
A;Molecule type: DNA
A;Residues: 1-193 <LIV>
A;Cross-references: EMBL:L42868; NID:g858735; PIDN:AAB37011.1; PID:g1695226
A;Experimental source: strain ZS7
C;Genetics:
A;Gene: ospC
C;Superfamily: Lyme disease spirochete surface protein C

Query Match   37.3%; Score 690.5; DB 2; Length 193;
Best Local Similarity   77.8%; Pred. No. 7.7e-27;
Matches 151; Conservative 10; Mismatches 30; Indels 3; Gaps 3;

Qy   3 CNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKAIG 62
            |||||||||||||||||||||||||||:||||||||||||||||||||||||||||:||| |
Db   1 CNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEI-AKAIG 59

Qy   63 KKIHQNNGLDTENNNGSLLAGAYAISTLIKQKLDGLK-NEGLKEKIDAARKCSETFTNK 121
            |||||||||||:|||||||||||||||||:|||||||||| :|||||||||||||:||:
Db   1 CNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEI-AKAIG 59

Qy   122 LKEKHTDLGKEGVTDADAKEAILKANGT-KTKGAEELGKLFESVEVLSKAAKEMLANSVK 180
            |:|| ||||||| ||| |||||||||||| ||:||||||||||||||||:||||||||:|
Db   120 LKDNHAQLGIGQVTDENAKKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANSVK 179

Qy   181 ELTSPVVAESPKKP 194
            ||||||||||||||
Db   180 ELTSPVVAESPKKP 193

RESULT 15
I40268
outer surface protein C precursor - Lyme disease spirochete
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 26-May-2000
C;Accession: I40268
R;Margolis, N.; Hogan, D.; Cieplak, W.
Gene 143, 105-110, 1994
A;Title: Homology between Borrelia burgdorferi OspC and members of the family of Borr
A;Reference number: I40268; MUID:94259285
A;Accession: I40268
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-211 <RES>
A;Cross-references: GB:L25413; NID:g495735; PIDN:AAA22956.1; PID:g495736
A;Genetics:
A;Gene: ospC
C;Superfamily: Lyme disease spirochete surface protein C

Query Match   37.1%; Score 686.5; DB 2; Length 211;
Best Local Similarity   73.8%; Pred. No. 1.3e-26;
Matches 144; Conservative 18; Mismatches 32; Indels 1; Gaps 1;

Qy   1 MACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKA 60
            ::|||||||||||||||||||||||||||:||||||||||||||||||||||||||:||| |
Db   17 ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDELAKKA 76

Qy   61 IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEG-LKEKIDAARKCSETFT 119

```
Db   77  IGKKIDQNNALGTLDNHNGSLLAGAYAISALITEKLSSIKDSGELKAEIEKAKKCCSESFT  136
            ||||| ||| | : :|||||||||||||||| || :||: | || ||: | :|||||:||
Qy  120  NKLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSV  179
            |: :|| | : ::|||: |||||| |||||: :|||||
Db  137  KKLSDNQAELGIENATDDNAKKAILKTHNAKDKGAEELVKLSESVAGLLKAAQAILANSV  196
            :||||||||||||||
Qy  180  KELTSPVVAESPKKP  194
            |||||||||||||||
Db  197  KELTSPVVAESPKKP  211
```

Search completed: March 18, 2002, 09:56:47
Job time: 370 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 10:11:45 ; Search time 39.62 Seconds
           (without alignments)
           348.880 Million cell updates/sec

Title: US-09-596-746A-30
Perfect score: 1850
Sequence: 1 MACNNSGKDGNTSANSADES..........AVENLAKAAKEMLANSVKEL 377

Scoring table: BLOSUM62
                  Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

al number of hits satisfying chosen parameters: 100059

imum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                    Maximum Match 100%
                    Listing first 45 summaries

Database :   SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	945	51.1	210	1	OSC1_BORBU	Q07337 borrelia bu
2	631	34.1	212	1	OSC2_BORBU	Q08137 borrelia bu
3	397.5	21.5	214	1	VM24_BORHE	P32778 borrelia he
4	397	21.5	215	1	VM03_BORHE	Q02448 borrelia he
5	174.5	9.4	1509	1	MYSN_ACACA	P05659 acanthamoeb
6	172	9.3	492	1	M5_STRPY	P02977 streptococc
7	164.5	8.9	1164	1	BAG_STRAG	P27951 streptococc
8	161.5	8.7	1790	1	USO1_YEAST	P25386 saccharomyc
9	160	8.6	1433	1	RESI_CHICK	O42184 gallus gall
10	159.5	8.6	1639	1	LMG1_DROME	P15215 drosophila
11	159.5	8.6	1957	1	YD86_SCHPO	Q10411 schizosacch
12	157	8.5	483	1	M6_STRPY	P08009 streptococc
13	151.5	8.2	2017	1	MYSN_DROME	Q09323 drosophila
14	150	8.1	1938	1	MYS_AEQIR	P24733 aequipecten
15	150	8.1	1966	1	MYSB_CAEEL	P02566 caenorhabdi
16	149	8.1	662	1	TLPA_BACSU	P39216 bacillus su
17	148	8.0	2116	1	MYS2_DICDI	P08799 dictyosteli
18	148	8.0	3110	1	LMA2_HUMAN	P24043 homo sapien
19	147.5	8.0	1013	1	SCA4_RICRH	Q09381 rickettsia
20	147.5	8.0	3210	1	CENP_HUMAN	P49454 homo sapien
21	147	7.9	564	1	M12_STRPY	P19401 streptococc
22	147	7.9	1230	1	SMC3_YEAST	P47037 saccharomyc
23	145	7.8	998	1	SCA4_RICAK	Q9aix9 rickettsia
24	143	7.7	1093	1	TMF1_HUMAN	P82094 homo sapien
25	142.5	7.7	2869	1	RBP1_PLAVB	Q00798 plasmodium
26	142	7.7	540	1	CH60_MYCPA	P42384 mycobacteri
27	142	7.7	3672	1	LML2_CAEEL	Q21313 caenorhabdi
28	141.5	7.6	875	1	ZIP1_YEAST	P31111 saccharomyc
29	141.5	7.6	1786	1	LMB1_MOUSE	P02469 mus musculu
30	141	7.6	1427	1	REST_HUMAN	P30622 homo sapien
31	140.5	7.6	1011	1	SCA_RICAF	Q9al83 rickettsia
32	140.5	7.6	1786	1	LMB1_HUMAN	P07942 homo sapien
33	140	7.6	1147	1	CGA1_HELPY	P80200 helicobacte
34	139.5	7.5	756	1	Y328_MYCGE	Q49419 mycoplasma
35	139.5	7.5	1935	1	MYSS_CYPCA	Q90339 cyprinus ca
36	139	7.5	1391	1	MST2_DROHY	Q08696 drosophila
37	139	7.5	2349	1	TPR_HUMAN	P12270 homo sapien
38	138.5	7.5	944	1	NUP1_YEAST	P32380 saccharomyc
39	138	7.5	724	1	HMMR_HUMAN	O75330 homo sapien
40	137.5	7.4	662	1	TLPB_BACSU	P39217 bacillus su
41	137	7.4	697	1	MFP1_LYCES	P93203 lycopersico
42	137	7.4	1500	1	SSP5_STPGN	P35748 streptococc
43	137	7.4	1972	1	MYHB_RABIT	P16952 oryctolagus
44	136.5	7.4	473	1	YCE8_BACSU	P40767 bacillus su
45	136.5	7.4	1189	1	YJH6_YEAST	P47035 saccharomyc

ALIGNMENTS

RESULT 1
OSC1_BORBU

ID   OSC1_BORBU   STANDARD;   PRT;   210 AA.
AC   Q07337;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   OUTER SURFACE PROTEIN C PRECURSOR (PC).
DE   OSPC OR BBB19.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OG   Plasmid lp54.
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   STRAIN=ATCC 35210 / B31;
RX   MEDLINE=93268136; PubMed=8098841;
RA   Jauris-Helpke S., Fuchs R., Motz M., Preac-Mursic V., Schwab E.,
RA   Will G., Wilske B.;
RT   "Genetic heterogenity of the genes coding for the outer surface
RT   protein C (OspC) and the flagellin of Borrelia burgdorferi.";
RL   Med. Microbiol. Immunol. 182:37-50(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   STRAIN=ATCC 35210 / B31;
RX   MEDLINE=93293332; PubMed=8478108;
RA   Wilske B., Preace-Mursic V., Jauris S., Pradel I., Soutschek E.,
RA   Schwab E., Wanner G.;
RT   "Immunological and molecular polymorphisms of OspC, an immunodominant
RT   major outer surface protein of Borrelia burgdorferi.";
RL   Infect. Immun. 61:2182-2191(1993).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   STRAIN=ATCC 35210 / B31;
RX   MEDLINE=94041630; PubMed=8225587;
RA   Padula S.J., Sampieri A., Dias F., Szczepanski A., Ryan R.W.;
RT   "Molecular characterization and expression of p23 (OspC) from a North
RT   American strain of Borrelia burgdorferi.";
RL   Infect. Immun. 61:5097-5105(1993).
RN   [4]
RP   SEQUENCE FROM N.A.
RX   STRAIN=ATCC 35210 / B31;
RX   MEDLINE=96025162; PubMed=7494039;
RA   Fukunaga M., Hamase A.;
RT   "Outer surface protein C gene sequence analysis of Borrelia
RT   burgdorferi sensu lato isolates from Japan.";
RL   J. Clin. Microbiol. 33:2415-2420(1995).
RN   [5]
RP   SEQUENCE FROM N.A.
RX   STRAIN=ATCC 35210 / B31;
RX   MEDLINE=9805943; PubMed=9403685;
RA   Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA   Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
RA   Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
RA   Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,
RA   van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,

---

RA   Utterback T., Watthey L., McDonald L., Artiach P., Bowman C.,
RA   Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
RA   Smith H.O., Venter J.C.;
RT   "Genomic sequence of a Lyme disease spirochaete, Borrelia
RT   burgdorferi.";
RL   Nature 390:580-586(1997).
-!- FUNCTION: NOT KNOWN; MAJOR IMMUNODOMINANT PROTEIN.
-!- SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID
     ANCHOR.

DR   EMBL; X69596; CAA49306.1; -.
DR   EMBL; U01894; AAA16058.1; -.
● DR   EMBL; D49497; BAA08457.1; -.
DR   EMBL; AE000792; AAC66329.1; -.
DR   TIGR; BB019; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
DR   PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KW   Outer membrane; Lipoprotein; Signal; Plasmid; Antigen;
KW   Complete proteome.
FT   SIGNAL   1   18
FT   CHAIN   19   210   OUTER SURFACE PROTEIN C.
FT   LIPID   19   19   N-ACYL DIGLYCERIDE (BY SIMILARITY).
SQ   SEQUENCE   210 AA; 22340 MW; 7AAFC978F91777BF CRC64;

Query Match
Best Local Similarity   97.9%;
Matches 190; Conservative   3; Pred. No. 3.9e-37; Mismatches   1; Indels   0; Gaps   0;
Score 945; DB 1; Length 210;

```
Qy    1  MACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKA   60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   17  ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA   76

Qy   61  IGKKIHQNNGLDTENNINGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN  120
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   77  IGKKIHQNNGLDTENNINGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN  136

Qy  121  KLKEKHFDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKRAAKEMLANSVK 180
            |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db  137  KLKEKHFDLGKEGVTDADAKEAILKINGTKTKGAEELGKLFESVEVLSKRAAKEMLANSVK 196

Qy  181  ELTSPVVAESPKKP  194
            ||||||||||||||
Db  197  ELTSPVVAESPKKP  210
```

RESULT 2
OSC2_BORBU
ID   OSC2_BORBU   STANDARD;   PRT;   212 AA.
AC   O08137;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   OUTER SURFACE PROTEIN C PRECURSOR (PC).
OS   OSPC.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
RC   STRAIN=PKO;
RX   MEDLINE=92219995; PubMed=1560779;
RA   Fuchs R., Jauris S., Lottspeich F., Preac-Mursic V., Wilske B.,
RA   Soutschek E.;
RT   "Molecular analysis and expression of a Borrelia burgdorferi gene
RT   encoding a 22 kDa protein (pc) in Escherichia coli.";
RL   Mol. Microbiol. 6:503-509(1992).
RN   [2]
RP   SEQUENCE OF 1-205 FROM N.A.
RC   STRAIN=DK26;
RX   MEDLINE=94075528; PubMed=8253951;
RA   Theisen M., Frederiksen B., Lebech A.M., Vuust J., Hansen K.;
RT   "Polymorphism in ospC gene of Borrelia burgdorferi and
RT   immunoreactivity of OspC protein: implications for taxonomy and for
RT   use of OspC protein as a diagnostic antigen.";
RL   J. Clin. Microbiol. 31:2570-2576(1993).
-!- FUNCTION: NOT KNOWN; MAJOR IMMUNODOMINANT PROTEIN.
-!- SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID
     ANCHOR.

DR   EMBL; X62162; CAA44093.1; -.
DR   EMBL; X73624; CAA52003.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
KW   Outer membrane; Lipoprotein; Signal; Plasmid; Antigen.
FT   SIGNAL   1   18   BY SIMILARITY.
FT   CHAIN   19   212   OUTER SURFACE PROTEIN C.
FT   LIPID   19   19   N-ACYL DIGLYCERIDE (BY SIMILARITY).
SQ   SEQUENCE   212 AA; 22499 MW; C206C231FBF2E7D4 CRC64;

Query Match
Best Local Similarity   69.9%;
Matches 137; Conservative 15; Pred. No. 7.6e-23; Mismatches 42; Indels   2; Gaps   2;
Score 631; DB 1; Length 212;

```
Qy    1  MACNNSGKDGNT-SANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAK   59
            ||||| ||| :|| |||||||||||||||||||| :||||||||||||||||:|||
Db   17  ISCNNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETLVLSIDELAKK   76

Qy   60  AIGKKIHQNNGLDTENNINGSLLAGAYAISTLIKQKLDGLKN-EGLKEKIDAAKKCSETF  118
            :||||| |  || |  |||| |||||||||||| :  ||| ||||||||||||| ||
Db   77  AIGQKIDNNNGLAALNNQNGSLLAGAYAISTLITEKLSKLKNLEELKTEIARAKKCSEEF  136

Qy  119  TNKLKEKHFDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKRAAKEMLANS 178
            ||||  ||| |:|||  ||| | | | ||||  | ||||||||||||||||||||||||
Db  137  TNKLKSGHADLGKQDATDDHAKAAILKHTHATTDKGAKEFKDLFESVEGLLKAAQVALTNS 196

Qy  179  VKELTSPVVAESPKKP  194
            ||||| ||||||||||
Db  197  VKELTSPVVAESPKKP  212
```

RESULT 3
VM24_BORHE
ID   VM24_BORHE   STANDARD;   PRT;   214 AA.
AC   P32778;
DT   01-OCT-1993 (Rel. 27, Created)
DT   01-OCT-1993 (Rel. 27, Last sequence update)
DT   01-FEB-1994 (Rel. 28, Last annotation update)
DE   VARIABLE MAJOR OUTER MEMBRANE LIPOPROTEIN 24 PRECURSOR.
DE   VMP24.
OS   Borrelia hermsii.
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=140;
RN   [1]
RP   SEQUENCE FROM N.A.

RESULT 4

```
VM03_BORHE
ID   VM03_BORHE     STANDARD;      PRT;   215 AA.
AC   Q02448;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   01-FEB-1994 (Rel. 28, Last annotation update)
DE   VARIABLE MAJOR OUTER MEMBRANE LIPOPROTEIN 3 PRECURSOR.
GN   VMP3.
OS   Borrelia hermsii.
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=140;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SSP. HS1 SEROTYPE 3;
RX   MEDLINE=93133110; PubMed=1484486;
RA   Restrepo B.I., Kitten T., Carter C.J., Infante D., Barbour A.G.;
RT   "Subtelomeric expression regions of Borrelia hermsii linear plasmids
RT   are highly polymorphic.";
RL   Mol. Microbiol. 6:3299-3311(1992).
CC   -!- FUNCTION: SERVES TO AVOID THE HOST IMMUNE RESPONSE BY CHANGING
CC       FROM ONE SURFACE EXPOSED VMP TO ANOTHER.
CC   -!- SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID
CC       ANCHOR.
CC   -!- SIMILARITY: STRONG, TO VMP3.
```

```
DR   EMBL; L04786; AAA22964.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
DR   PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KW   Outer membrane; Lipoprotein; Signal; Plasmid.
FT   SIGNAL        1     18
FT   CHAIN        19    214       VARIABLE MAJOR OUTER MEMBRANE
FT                                LIPOPROTEIN 24.
FT   LIPID        19     19       N-ACYL DIGLYCERIDE (PROBABLE).
SQ   SEQUENCE   214 AA;  22541 MW;  F1583F510246F7C7 CRC64;

Query Match             21.5%;  Score 397.5;  DB 1;  Length 214;
Best Local Similarity   44.8%;  Pred. No. 3.2e-12;
Matches   91;  Conservative   37;  Mismatches   58;  Indels   17;  Gaps   5;

Qy     1 MACNNSGKDGNTSANSADESVKGP----NLTEINKKITDSNAVLLAVKEVEALLSSIDEI 56
           : ||| :|| | |: ||||: | |:  |||||||:|||:|   :|:
Db    17 MSCNNGGPE----LKSDEVAKSDGTVLDLAKVSKKIKEASAPAASVKEVETLVKSVDEL 71

Qy    57 AAKAIGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKL------DGLKNEGLKEKIDA 110
          |||||||| :| ||||| ::|||:|| : : :| ||           :| ||: : |
Db    72 -AKAIGKKIKNDGGLDTEAGQNGSLIAGVHSVSVSAVKIKVGALETTSGISNE-LKTKITE 129

Qy   111 AKKCSETFTNKLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKA 170
          .:. :| : ||| || |:|||: | :  || |||| ||||||: || : | : :| |.|
Db   130 VKSKAEAFLNKLKDGHTELGKKDASDDTKKAIKKDNSDKTKGASELEALNTAVDALLKA 189

Qy   171 AKEMLANSVKELTSPVVAESPKK 193
           : :  |:||||:|| |||:|
Db   190 AEGEVEAAIKELTAPVKAEKPSQ 212
```

RESULT 5

```
MYSN_ACACA
ID   MYSN_ACACA     STANDARD;      PRT;   1509 AA.
AC   P05659;
DT   01-NOV-1988 (Rel. 09, Created)
DT   01-NOV-1988 (Rel. 09, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   MYOSIN II HEAVY CHAIN, NON MUSCLE.
OS   Acanthamoeba castellanii (Amoeba).
OC   Eukaryota; Acanthamoebidae; Acanthamoeba.
OX   NCBI_TaxID=5755;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87308395; PubMed=3040773;
RA   Hammer J.A. III, Bowers B., Paterson B.M., Korn E.D.;
RT   "Complete nucleotide sequence and deduced polypeptide sequence of a
RT   nonmuscle myosin heavy chain gene from Acanthamoeba: evidence of a
RT   hinge in the rodlike tail.";
RL   J. Cell Biol. 105:913-925(1987).
CC   -!- FUNCTION: MYOSIN IS A PROTEIN THAT BINDS TO F-ACTIN & HAS ATPASE
CC       ACTIVITY THAT IS ACTIVATED BY F-ACTIN.
CC   -!- SUBUNIT: MYOSIN II HEAVY CHAIN IS TWO-HEADED. IT SELF-ASSEMBLES
CC       INTO FILAMENTS. HEXAMER OF 2 HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI
```

The corresponding database entry (VMP24) and alignment:

```
RT   are highly polymorphic.";
RL   Mol. Microbiol. 6:3299-3311(1992).
CC   -!- FUNCTION: SERVES TO AVOID THE HOST IMMUNE RESPONSE BY CHANGING
CC       FROM ONE SURFACE EXPOSED VMP TO ANOTHER.
CC   -!- SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID
CC       ANCHOR.
CC   -!- SIMILARITY: STRONG, TO VMP24.
```

```
DR   EMBL; L04789; AAA22967.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
DR   PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KW   Outer membrane; Lipoprotein; Signal; Plasmid.
FT   SIGNAL        1     18
FT   CHAIN        19    215       VARIABLE MAJOR OUTER MEMBRANE
FT                                LIPOPROTEIN 3.
FT   LIPID        19     19       N-ACYL DIGLYCERIDE (PROBABLE).
SQ   SEQUENCE   215 AA;  23139 MW;  684C74D35F87C771 CRC64;

Query Match             21.5%;  Score 397;  DB 1;  Length 215;
Best Local Similarity   45.3%;  Pred. No. 3.4e-12;
Matches   92;  Conservative   35;  Mismatches   60;  Indels   16;  Gaps   5;

Qy     1 MACNNSGKDGNTSANSADESVKGP----NLTEINKKITDSNAVLLAVKEVEALLSSIDEI 56
Db    17 MSCNNGGPE----LKSDEVAKSDGTVLDLAKISKKIKDASDFAASVKEVHTLVKSIDEL 71

Qy    57 AAKAIGKKI-HQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLK-----NEGLKEKIDA 110
Db    72 -AKAIGKKIKNDSNFEDENDHNGSLIAGVFQVILTVKAKLTSLEQIGISDELKTEVGM 130

Qy   111 AKKCSETFTNKLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKA 170
Db   131 VKKESEAFVTQVKSKHTDLAKEGVTDAHAKSAILVTDGTKDKGAAELIKLNTAIDELLKA 190

Qy   171 AKEMLANSVKELTSPVVAESPKK 193
Db   191 ANDAVETVIKELTASVKAEKPSQ 213
```

```
CC        LIGHT CHAIN SUBUNITS (MLC) AND 2 REGULATORY LIGHT CHAIN SUBUNITS
CC        (MLC-2).
-!-  DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CC        CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CC        CHARACTERISTIC OF ALPHA-HELICAL STRUCTURES. THIS REGION IS
CC        INTERRUPTED BY A HINGE AND JOINED BY A NONHELICAL TAILPIECE WHERE
CC        THE REGULATORY PHOSPHORYLATION SITES RESIDE.
-!-  MISCELLANEOUS: THE HINGE REGION MAY PLAY A KEY ROLE IN MEDIATING
CC        THE EFFECT OF HEAVY CHAIN PHOSPHORYLATION ON ENZYMATIC ACTIVITY.
-!-  SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
-!-  SIMILARITY: CONTAINS 2 IQ DOMAINS.
```

```
DR   EMBL; Y00624; CAA68663.1; -.
DR   PIR; A27224; A27224.
     HSSP; P08799; 1MND.
DR   InterPro; IPR000048; IQ.
DR   InterPro; IPR001609; myosin_head.
DR   Pfam; PF00612; IQ; 2.
DR   Pfam; PF00063; myosin_head; 1.
DR   PRINTS; PR00193; MYOSINHEAVY.
DR   ProDom; PD000355; myosin_head; 1.
DR   SMART; SM00015; IQ; 1.
DR   SMART; SM00242; MYSc; 1.
DR   PROSITE; PS50096; IQ; 1.
KW   Myosin; Coiled coil; Actin-binding; ATP-binding; Calmodulin-binding;
KW   Methylation; Alkylation; Phosphorylation; Multigene family.
FT   DOMAIN      1     789    MYOSIN HEAD-LIKE.
FT   DOMAIN    790     819    IQ.
FT   DOMAIN    848    1509    COILED COIL (POTENTIAL).
FT   DOMAIN    848    1226    ALPHA-HELICAL TAILPIECE (S2).
FT   DOMAIN   1227    1252    HINGE.
FT   DOMAIN   1253    1509    LIGHT MEROMYOSIN (LMM).
FT   DOMAIN   1253    1482    ALPHA-HELICAL TAILPIECE (LMM).
FT   DOMAIN   1483    1509    NONHELICAL TAILPIECE (LMM).
FT   NP_BIND   182     189    ATP.
FT   DOMAIN    660     682    ACTIN-BINDING.
FT   DOMAIN    766     780    ACTIN-BINDING.
FT   MOD_RES   133     133    METHYLATION (TRI-) (POTENTIAL).
FT   MOD_RES   700     700    ALKYLATION (SH-1) (POTENTIAL).
FT   MOD_RES  1489    1489    PHOSPHORYLATION.
FT   MOD_RES  1494    1494    PHOSPHORYLATION.
FT   MOD_RES  1499    1499    PHOSPHORYLATION.
SQ   SEQUENCE  1509 AA;  171201 MW;  2CE49BE51173D17E CRC64;
```

```
Query Match   9.4%;  Best Local Similarity  24.7%;
Matches  98;  Conservative  73;  Mismatches 159;  Indels  67;  Gaps  19;
             Score 174.5;  DB 1;  Length 1509;  Pred. No. 0.38;

Qy    28  EINKKITDSNAVLLAVKEVEA----LLSSIDEIAKAIGKKIHQ-NNGLDTENNHNGSL   81
          :|:|  ||||   |:|  | :  |  ||  |:::|:   |||| |:||   :||| ||||
Db   934  ELQEETSASNDILEQRKLEAEKGELKASLEE--EERNRKALQEAKTKVESERNELQDK  990

Qy    82  LAGAYAISTLIKQKLDGLKNEGLKEKIDA----AKKCSEFPTNKLKEKHVDLGKEGVTDA-  137
          ||||   |    |||  | ||| |:||||    ||| |::| ||||| |:|:|  |: |
Db   991  YEDEAAAHDSLKKEEDLSRE-LRETKDALADAENISETLRSKL--KNTERGADDVRNEL  1047

Qy   138  -DAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKELTSPV---------   186
           ||:: | |  :|:| | :||| ||  | |||:|||  |:|:| |::| |
Db  1048  DDVTATKLQLEKTKKSLEEBLAQTRAQLEE-EKSGKEAASSKAKOLGGOLEDARSEVDSL  1106

Qy   187  ----VAESPKKPSMVNNSG----KDGNTSANSADESVKG--PNLTEISKKIT----ES   230
           |:|        || |     |:|     |:|| ::| |   |:| |:||:|   ::|
Db  1107  KSKLSAAEKSLKTAKDQNRDLDEQLEDERTVRANVDKQKKALEAKLTELEDQVTALDGQK  1166
```

```
RESULT   6
ID   M5_STRPY        STANDARD;       PRT;    492 AA.
AC   P02977;
DT   21-JUL-1986 (Rel. 01, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   M PROTEIN, SEROTYPE 5 PRECURSOR.
GN   EMM5 OR SMP5.
OS   Streptococcus pyogenes.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC   Streptococcus.
OX   NCBI_TaxID=1314;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88168881; PubMed=3281944;
RA   Miller L., Gray L., Beachey E., Kehoe M.;
RT   "Antigenic variation among group A streptococcal M proteins.
RT   Nucleotide sequence of the serotype 5 M protein gene and its
RT   relationship with genes encoding types 6 and 24 M proteins.";
RL   J. Biol. Chem. 263:5668-5673(1988).
RN   [2]
RP   SEQUENCE OF 43-212 AND 238-250.
RX   MEDLINE=84162039; PubMed=6368549;
RA   Manjula B.N., Acharya A.S., Mische S.M., Fairwell T., Fischetti V.A.;
RT   "The complete amino acid sequence of a biologically active
RT   197-residue fragment of M protein isolated from type 5 group A
RT   streptococci.";
RL   J. Biol. Chem. 259:3686-3693(1984).
CC   -!- FUNCTION: THIS PROTEIN IS ONE OF THE DIFFERENT ANTIGENIC SEROTYPES
CC       OF PROTEIN M. PROTEIN M IS CLOSELY ASSOCIATED WITH VIRULENCE OF
CC       THE BACTERIUM AND CAN RENDER THE ORGANISM RESISTANT TO
CC       PHAGOCYTOSIS.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. CELL WALL.
CC   -!- SIMILARITY: TO OTHER M PROTEINS.
CC   -!- SIMILARITY: TO OTHER STREPTOCOCCAL AND STAPHYLOCOCCAL PROTEINS
CC       IN THE REGION OF THE MEMBRANE ANCHOR.
```

```
DR   EMBL; M20374; AAA26976.1; -.
DR   PIR; A03501; MMSOMP.
DR   PIR; A28616; A28616.
DR   InterPro; IPR001899; Gram_pos_anchor.
DR   InterPro; IPR003345; M_repeat.
DR   Pfam; PF00746; Gram_pos_anchor; 1.
DR   Pfam; PF02370; M; 9.
DR   PRINTS; PR00015; GPOSANCHOR.
DR   PROSITE; PS00343; GRAM_POS_ANCHORING; 1.
KW   Virulence; Phagocytosis; Cell wall; Duplication; Repeat; Antigen;
KW   Transmembrane; Coiled coil; Signal.
FT   SIGNAL      1      42
FT   CHAIN      43     492    M PROTEIN, SEROTYPE 5.
FT   DOMAIN     43     466    EXTRACELLULAR (POTENTIAL).

Qy   231  NAVVLAVKEIETLLASIDELATKAIGKKIQONGGIA--VEAGHNGTLLAGAYTISKLITQ   288
          | ||  |:||  |||   || | :| | | |: ||| | ||:  |:| |  |:|:||||
Db  1167  NA---AAAQARTLKTQVDE--TK---RRLEEAFASAARLEKERKNALDEVAQLTADLDAE  1218

Qy   289  KLDGLKNSEKLKEKI------------ENAKK----CSEDFTKKLEGEHAQLGIENVTDENAKAI   338
          ||  |: :|  |:|              ||:|    |::|   ||||   |:|:|    |:||
Db  1219  RDSGAQQRRKLNTRISELQSELENAPKTGGASSEEVKRLEGELRLEEILTAQEARAA-   1277

Qy   339  LITDAAKDKGAAELEKLFKAVENLAKAAKEMLANSVK   375
          |   |  |||||| || |  |  ||| |  |:| |||
Db  1278  --AEKNLDKRANLELEELRQEADDAARDNDKLVKDNRK 1312
```

Left column

```
FT   TRANSMEM     467    486       MEMBRANE ANCHOR.
FT   DOMAIN       487    492       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       421    457       GLY/PRO-RICH (CELL WALL-SPANNING).
FT   DOMAIN       458    463       CONSERVED IN GRAM-POSITIVE COCCI SURFACE
FT                                 PROTEINS.
FT   DOMAIN        69    103       5 X 7 AA TANDEM REPEATS OF L-K-T-K-N-E-G.
FT   REPEAT        69     75       1.
FT   REPEAT        76     82       2.
FT   REPEAT        83     89       3.
FT   REPEAT        90     96       4.
FT   REPEAT        97    103       5.
FT   CONFLICT      43     43       A -> T (IN REF. 2).
FT   CONFLICT      50     50       N -> S (IN REF. 2).
FT   CONFLICT     102    102       K -> L (IN REF. 2).
FT   CONFLICT     208    208       I -> L (IN REF. 2).
FT   CONFLICT     102    102       SNLERKTAELTSEK (IN REF. 2).
FT   CONFLICT     208    208       SNLERKTAELTSEK (IN REF. 2).
SQ   SEQUENCE     492 AA;  55085 MW;  87379B6CBD55E27 CRC64;
```

Query Match   9.3%; Score 172; DB 1; Length 492;
Best Local Similarity   24.9%; Pred. No. 0.14;
Matches 107; Conservative 54; Mismatches 150; Indels 118; Gaps 18;

```
Qy    31  KKITDSNAVLLAV---------KEVEALLS-SIDEIAAKAIGKKIHQNNGLDTENHH   77
Db    16  KKGTASVAVALSVLGAGLVVNTNEVSAAVTRGTINDPQRAKEALDKYELENHDLKTKNEG  75

Qy    78  NGSLLAGAYAISTLIKQKLDGLK---NEGLK----EKIDAAKKCSETFNKLKEKHTDLGKE  132
Db    76  --------LKTENEGLKTENEGLKTENEGLKTEKKEHEAENDKLKQQRDTLSTQ  121

Qy   133  GVT-DADAK-------EAILKANGTKTI----KGAEELGKLFESVEVLSKAAKEMIANSVKE  181
Db   122  KETLEREVQNTQYNNEFLKIKNGDLTKELNKTROELANKQQESKENEKALNELLEKTVKD  181

Qy   182  LTSPVVAESPKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIE  241
Db   182  --KIAKEQENKETI--GTLKKILDETVKDKIKIAKEQENKET----------IG  219

Qy   242  TLLASIDELATKAIGK---KIQON-GGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEK  298
Db   220  TLKKILDETVKDKLAKEQKSKONIGALKQELAKKDEANKISDASRKGLRRDLDASREAKK  279

Qy   299  LKE----KIENAKKCSE----------------DFTKKLEGEHAQLGIENVTDENAKKAI-L  339
Db   280  QLEAEHQKLEEQNKISEASRKGLRRDLDASREAKKQLEAEQQKLEEQNKISEASRKGLRR  339

Qy   340  ITDAAKD------------------KGAAELEKLFKAVENL---------------AKAAKE  368
Db   340  DLDASREAKKQVEKALEEANSKLAALEKLNKELEESKKLTEKEKAELQAKLEAEAKALKE  399

Qy   369  MLANSVKEL  377
Db   400  QLAKQAEEL 408
```

RESULT 7
BAG_STRAG

Right column

CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. CELL WALL.
CC   -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE DOMAIN.
CC   -!- SIMILARITY: TO OTHER STREPTOCOCCAL AND STAPHYLOCOCCAL PROTEINS
CC   IN THE REGION OF THE MEMBRANE ANCHOR.

```
DR   EMBL; X59771; CAA42442.1; -.
DR   PIR; S15330; FCSOAG.
DR   InterPro; IPR001899; Gram_pos_anchor.
DR   InterPro; IPR003599; Ig.
DR   Pfam; PF00746; Gram_pos_anchor; 1.
DR   SMART; SM00409; IG; 1.
DR   PROSITE; PS00343; GRAM_POS_ANCHORING; 1.
KW   Cell wall; Transmembrane; Receptor; Repeat; Signal;
KW   Immunoglobulin domain.
FT   SIGNAL         1     37
FT   CHAIN         38   1164       IGA FC RECEPTOR.
FT   DOMAIN        38   1131       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    1141   1159       MEMBRANE ANCHOR (POTENTIAL).
FT   DOMAIN      1160   1164       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       434    534       IG-LIKE DOMAIN.
FT   DOMAIN       199    438       IG-BINDING (POTENTIAL).
FT   DOMAIN       439    826       IG-BINDING (POTENTIAL).
FT   DOMAIN       827    945       PRO-RICH REPEATS.
FT   DOMAIN      1131   1137       CONSERVED IN GRAM-POSITIVE COCCI SURFACE
FT                                 PROTEINS.
SQ   SEQUENCE    1164 AA;  131051 MW;  65DE94AF720A5474 CRC64;
```

Query Match   8.9%; Score 164.5; DB 1; Length 1164;
Best Local Similarity   25.1%; Pred. No. 0.81;
Matches 95; Conservative 55; Mismatches 147; Indels 81; Gaps 18;

```
Qy     6  SGKDGNTSANSADESVKGPNLTEINRKKITDSNAVLLAVK------EVEALLSSIDEIAAKAIG   62
Db   128  SHEQKNEFIKIDETNDSDALLELENQFNETNR-LLHIKQHEEEVERDKKAKQQKTLKQSD  186

Qy    63  KKIHQNNGLDTENHHGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNKL  122
Db   187  TKVDLSN-IDKELNHGKS-----QVEKMAEQK--GITNED--------------KDSM  222

Qy   123  KEKHTDLGKEGVTDADAKEAILKANGTKTKGAEEGLKLFESVEVLSKAAKEMIANSVKEL  182
Db   223  LKKIEDIRKQ-AQQADKKE-----DAEVKVREELGKLFSS----TKAG---------L  261

Qy   183  TSPVVAESPKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVL------  235
Db   262  DQEIQEHVKETSSEENTQKVDEHVANSL-QNLAQKSLEELDKATTNEQATQVKNQFLEN  320

Qy   236  --AVKEIETLLASIDELATKAIGKKIQQNGGLAVEAGHNG-----TLLAGAYTISKLITQ  288
Db   321  AQKLKEIQPLIKETNVKLYKAMSESLEQ---VEKELKHNSEANLEDLVAKSKEIVREYEG  377

Qy   289  KLDGLKNSEKLKEKIENA------KRCSEDFTKKLEGEHAQLGIENVTDENAKKAILITDAA  344
Db   378  KLNQSKNLPELKQLEEEAHSKLKLKQVVEDFRKKFKTS-----EQVTPKKRVKRDL---AA  428
```

Qy   345 KDKGAAELEKLFKAVENL 362
            ::|:| :||| :|||
Db   429 NENNQQKIE-LTVSPENI 445

RESULT 8
ID   USOl_YEAST   STANDARD;   PRT;   1790 AA.
AC   P25386;
DT   01-MAY-1992 (Rel. 22, Created)
DT   01-MAY-1992 (Rel. 22, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   INTRACELLULAR PROTEIN TRANSPORT PROTEIN USOl.
GN   USOl OR INT1 OR YDL058W
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
     NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=X2180-1A;
RX   MEDLINE=91185402; PubMed=2010462;
RA   Nakajima H., Hirata A., Ogawa Y., Yonehara T., Yoda K.,
RA   Yamasaki M.;
RT   "A cytoskeleton-related gene, usol, is required for intracellular
RT   protein transport in Saccharomyces cerevisiae.";
RL   J. Cell Biol. 113:245-260(1991).
RN   [2]
RP   SEQUENCE OF 782-1790 FROM N.A.
RA   Hostetter M.K., Herman D.J., Bendel C.M., McClellan M., Tao N.,
RA   Kendrick K.E.;
RL   Submitted (FEB-1993) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 1-8 FROM N.A.
RA   Bai Y., Symington L.S.;
RL   Submitted (MAY-1996) to the EMBL/GenBank/DDBJ databases.
-!- FUNCTION: REQUIRED FOR PROTEIN TRANSPORT FROM THE ER TO THE GOLGI
     COMPLEX.
-!- SUBCELLULAR LOCATION: CYTOPLASMIC. ASSOCIATED WITH INTRACELLULAR
     MEMBRANES. PROBABLY PRESENT ON VESICLES OPERATIONAL BETWEEN THE
     ER AND THE GOLGI COMPLEX.
-!- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, COMPOSED
     OF AN HEPTAPEPTIDE REPEAT PATTERN CHARACTERISTIC OF ALPHA-HELICAL
     COILED COILS. MAY FORM FILAMENTOUS STRUCTURES IN THE CELL.
-!- SIMILARITY: BELONGS TO THE VDP/USOl/YBL047C FAMILY.

DR   EMBL; X54378; CAA38253.1; -.
DR   EMBL; L03188; AAB00143.1; -.
DR   EMBL; U53668; AAB66659.1; -.
DR   PIR; A38455; A38455.
DR   HSSP; P80220; 1DIP.
     SGD; S000216; USOl.
     InterPro; IPR002017; Spectrin.
KW   Transport; Protein transport; Golgi stack; Cytoskeleton; Coiled coil.
     DOMAIN   1   724   GLOBULAR HEAD.
     DOMAIN   725   1790   COILED COIL (POTENTIAL).
     DOMAIN   465   487   CHARGED (HYPER-HYDROPHILIC).
     DOMAIN   991   1790   DISPENSABLE FOR THE PROTEIN FUNCTION.
     DOMAIN   1172   1786   ASP/GLU-RICH (ACIDIC).
     CONFLICT   847   847   G -> E (IN REF. 2).
     CONFLICT   924   924   E -> K (IN REF. 2).
     CONFLICT   1253   1253   V -> I (IN REF. 2).
     CONFLICT   1319   1319   I -> V (IN REF. 2).
     CONFLICT   1461   1461   N -> S (IN REF. 2).
FT   CONFLICT   1581   1581   G -> S (IN REF. 2).
FT   CONFLICT   1600   1600   I -> V (IN REF. 2).
FT   CONFLICT   1661   1661   R -> S (IN REF. 2).
FT   CONFLICT   1772   1772   D -> DEEDDEE (IN REF. 2).
SQ   SEQUENCE   1790 AA; 206424 MW; 6CE2B216E9FD4818 CRC64;

Query Match 8.7%; Score 161.5; DB 1; Length 1790;
Best Local Similarity 21.7%; Pred. No. 1.8;
Matches 99; Conservative 72; Mismatches 178; Indels 107; Gaps 16;

Qy   3 CNNSGKDG---------------NTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSS 52
          ||| : |
Db   929 CNNLSKEKEHISKELVEYKSRFQSHDNLV--AKLTEKLKSLANNYKDMQA--ENESLIKA 984

Qy   53 IDEIAAKAIGKKIHQNNGLD--TENHNGSLLAGAY------AISTL------IKQK 95
          :|:| :: :| |: |||| ::|| :|||:
Db   985 VEESKNESSIQLSNLQNKIDSMSQEKENFQIERGSIEKNIEQLKKTISDLEQTKEEIISK 1044

Qy   96 LDGLKNEG------LKEKIDAAKKCSETFTNKLKE---------KHTDLGKEGVTD 136
          ||||: ||||| : :: :| ::|
Db 1045 SDSSKDEYESQISLLKEKLETFATTANDENVNKISELTKTREELEAELAAYKNLKNELETK 1104

Qy   137 ADAKEAILK------------ANGTKTK------GAEELGKLFESVEVLSKAA 171
          | |:|:|: |
Db 1105 LETSEKALKEVKENEEHLKEEKIQLEKEATETKQQLNSLRANLESLEKEHEDLAAQLKKY 1164

Qy   172 KEMLANSVKELTSPV--VAESPKKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITE 229
          |: : :   : ::: : |:|
Db 1165 EEQIANKERQYNEEISQLNDEITSTQQENESIKKNDELEGEVKRAMKSTSEEQSNLKKSE 1224

Qy   230 SNAVLAVKKIE--------TLLASIDELATKAIGKKIQQNGGLAVEAGHNGTLLAGAYT 281
          | : :|| :   : | :
Db 1225 IDALNLQIKELKKKNETNEASLLESIKSVESETY--KIKE---LQDECNFK-------- 1270

Qy   282 ISKLITQKLDGLKJNSEKLKEKIENAKKCSEDFTKKLEGEHAQLGIENVTDENAKKAILIT 341
          | :| |||| ||| : : | ::||
Db 1271 -EKEVSELEDKIKASEDKNSKYLELQKESEKIKEELDAKTTELKIQLEKITINLSK----- 1324

Qy   342 DAAKDKGAAELEKLFKAVENLAKAAKEMLANSVKEL 377
          :| :| ||| || |:| |
Db 1325 --AKEKSESELSRLKKTSSEERNAEQLEKLKNEI 1358

RESULT 9
ID   REST_CHICK   STANDARD;   PRT;   1433 AA.
AC   041184; 042228; O57563; 057564;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   RESTIN (CYTOPLASMIC LINKER PROTEIN-170) (CLIP-170).
GN   RSN.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
     NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98137792; PubMed=9469933;
RA   Griparic L., Volosky J.M., Keller T.C. III;
RT   "Cloning and expression of chicken CLIP-170 and restin isoforms.";
RL   Gene 206:195-208(1998).
RN   [2]
RP   SEQUENCE OF 17-1139 FROM N.A. (CLIP-170(11) AND CLIP-170(11+35)).
RC   TISSUE=Pectoralis muscle;
RA   Griparic L., Keller T.C. III;
RT   "Identification and expression of two novel CLIP-170/Restin isoforms
RT   expressed predominantly in muscle.";
RL   Submitted (FEB-1998) to the EMBL/GenBank/DDBJ databases.
-!- FUNCTION: SEEMS TO BE A INTERMEDIATE FILAMENT ASSOCIATED PROTEIN
     THAT LINKS ENDOCYTIC VESICLES TO MICROTUBULES (BY SIMILARITY).

```
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC, ASSOCIATED WITH THE
CC       CYTOSKELETON (BY SIMILARITY).
CC   -!- ALTERNATIVE PRODUCTS: AT LEAST 4 ISOFORMS ARE PRODUCED BY
CC       ALTERNATIVE SPLICING.
CC   -!- SIMILARITY: CONTAINS 2 CAP-GLY DOMAINS.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AF014012; AAC60344.1; -.
DR   EMBL; AF020764; AAC60345.1; -.
DR   EMBL; AF045650; AAC03547.1; -.
DR   EMBL; AF045651; AAC03548.1; -.
DR   InterPro; IPR000938; CAP-Gly.
DR   InterPro; IPR001878; ZnF_CCHC.
DR   Pfam; PF01302; CAP_Gly; 2.
DR   SMART; SM00343; ZnF_C2HC; 1.
DR   PROSITE; PS00845; CAP_Gly; 2.
KW   Cytoskeleton; Microtubules; Coiled coil; Alternative splicing.
FT   DOMAIN        79    121       CAP-GLY 1.
FT   DOMAIN       144    207       SER-RICH.
FT   DOMAIN       235    277       CAP-GLY 2.
FT   DOMAIN       305    332       SER-RICH.
FT   DOMAIN       351   1353       COILED COIL (POTENTIAL).
FT   DOMAIN      1414   1427       CCHC-BOX.
FT   DOMAIN       458    492       MISSING (IN SHORT ISOFORM).
FT   VARSPLIC     458    492       TOTKLEHARIKELEQSLLFEKTKADRLQRELEDTR ->
FT   VARSPLIC     803    803       RKRQISEDPEN (IN ISOFORM CLIP-170(11)).
FT   VARSPLIC     458    492       S -> GGSKKVS (IN ISOFORM CLIP-170(11)).
FT   VARSPLIC     803    803       T -> RKRQISEDPENT (IN ISOFORM CLIP-
FT                                 170(11+35)).
FT   CONFLICT     309    309       K -> R (IN REF. 2; AAC03547).
FT   CONFLICT     440    440       K -> R (IN REF. 2; AAC03547).
FT   CONFLICT     309    309       E -> V (IN REF. 2; AAC03548).
SQ   SEQUENCE  1433 AA;  161026 MW;  5631C868B3498E23 CRC64;

Query Match             8.6%;  Score 160;  DB 1;  Length 1433;
Best Local Similarity  25.9%;  Pred No.: 1.6;
Matches   96;  Conservative  58;  Mismatches 158;  Indels  58;  Gaps  17;

Qy   28  EINKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGKKIHQNGLDTENNHNGSLLAGAYA  87
          :||:|| : ||||  ||
Db  524  ESSKHIDVDTLSLLQEI----SSLQEKMA-AAGKE-HQREMSSLKEFESSEEALRKE  577

Qy   88  ISTL-IKQKLDGLKNEGLKEKIDAAKK----CSETFTNKLKE--KHTDLGKE-------  132
          : |: | :|| |  :  |:  |:
Db  578  IKTLSASNERMGKENESLKTKLDHANKENSDVIELWKSKLESAIASHQQAMEELKVSFNK  637

Qy  133  GVTDADAKEAILKANGTKTKGAEELGKLPESVEVLSKAAKEMLANSVKELTSPVVAESPK  192
          || :  |  | |  |
Db  638  GVGAQTAEFAELKTOMEKVKLDYE--NEMSNLKLKQENEKSQHLKEIEAIKAKLLEVTEE  695

Qy  193  KPPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDELAT  252
          |  :  :|:  |  :
Db  696  KEQTLEN-------LKARLESVEDQHLVEMEDTLNKLQEAEIKVKELDVLQAKCNE-QT  746

Qy  253  KAIGKKIQQNGLAVEAGHNGTL----LAGAYITSKLITQKL-DGLKNSEKLKEKIENAK  307
          |  : : :  |  :    :|
Db  747  KLIGSLTQQ-----IRASEEKLDLAALQKANSEGKLEIOKLSEQLQAAEKQIONLETEK  801

Qy  308  KCSEDFTKLEGEHAQLGIENVTDENAKKAILITDAAKDKGAAEL-----EKLFKAV--E  360
          | |   :|: |:| :|   :   :| :  :: : |
Db  802  --VSNLTKELQGKQKL------LDLEKNLSAVNQVKDSLEKELQLLKEKFTSAVDGAE  852

Qy  361  NLAKAAKEML  370
          | |: :::|:
Db  853  NAQRAMQETI  862
```

RESULT 10

```
ID   LMG1_DROME     STANDARD;      PRT;   1639 AA.
AC   P15215; Q24373; Q9VT18;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   LAMININ GAMMA-1 CHAIN PRECURSOR (LAMININ B2 CHAIN).
GN   LANB2 OR LAMC1 OR LAMG1 OR CG3322.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=OREGON-R;
RX   MEDLINE=89109164; PubMed=2912972;
RA   Chi H.-C., Hui C.-F.;
RT   "Primary structure of the Drosophila laminin B2 chain and comparison
RT   with human, mouse, and Drosophila laminin B1 and B2 chains.";
RL   J. Biol. Chem. 264:1543-1550(1989).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=OREGON-R;
RX   MEDLINE=90377237; PubMed=2808533;
RA   Montell D.J., Goodman C.S.;
RT   "Drosophila laminin: sequence of B2 subunit and expression of all
RT   three subunits during embryogenesis.";
RL   J. Cell Biol. 109:2441-2453(1989).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Chi H.-C., Juminaga D., Wang S.Y., Hui C.-F.;
RT   "Structure of the Drosophila gene for the laminin B2 chain.";
RL   DNA Cell Biol. 10:451-466(1991).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
```

```
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [5]
RP   SEQUENCE OF 344-1639 FROM N.A.
RC   STRAIN=OREGON-R;
RX   MEDLINE=88303364; PubMed=3405777;
RA   Chi H.-C., Hui C.-F.;
RT   "cDNA and amino acid sequences of Drosophila laminin B2 chain.";
RL   Nucleic Acids Res. 16:7205-7205(1988).
```

```
CC   -!- FUNCTION: BINDING TO CELLS VIA A HIGH AFFINITY RECEPTOR, LAMININ
CC       IS THOUGHT TO MEDIATE THE ATTACHMENT, MIGRATION, & ORGANIZATION OF
CC       CELLS INTO TISSUES DURING EMBRYONIC DEVELOPMENT BY INTERACTING
CC       WITH OTHER EXTRACELLULAR MATRIX COMPONENTS.
CC   -!- SUBUNIT: LAMININ IS A COMPLEX GLYCOPROTEIN, CONSISTING OF THREE
CC       DIFFERENT POLYPEPTIDE CHAINS (ALPHA, BETA, GAMMA), WHICH ARE BOUND
CC       TO EACH OTHER BY DISULFIDE BONDS INTO A CROSS-SHAPED MOLECULE
CC       COMPRISING ONE LONG & THREE SHORT ARMS WITH GLOBULES AT EACH END.
CC   -!- SUBCELLULAR LOCATION: BASEMENT MEMBRANES (MAJOR COMPONENT).
CC   -!- DOMAIN: THE ALPHA-HELICAL DOMAINS I AND II ARE THOUGHT TO INTERACT
CC       WITH OTHER LAMININ CHAINS TO FORM A COILED COIL STRUCTURE.
CC   -!- DOMAIN: DOMAINS VI AND IV ARE GLOBULAR.
CC   -!- SIMILARITY: CONTAINS 1 LAMININ N-TERMINAL DOMAIN (DOMAIN VI).
CC   -!- SIMILARITY: CONTAINS 11 LAMININ EGF-LIKE DOMAINS.
CC   -!- SIMILARITY: CONTAINS 1 LAMININ DOMAIN IV.
```

```
DR   EMBL; M58417; AAA28665.1; -.
DR   EMBL; M25063; AAA28664.1; -.
DR   EMBL; X07806; AAF50238.1; -.
DR   EMBL; X07806; CAA30665.1; -.
DR   PIR; A31483; MMFFB2.
DR   HSSP; P02468; 1TLE.
DR   FlyBase; FBgn0002528; LanB2.
DR   InterPro; IPR000561; EGF-like.
DR   InterPro; IPR001886; LamNT.
DR   InterPro; IPR000034; Laminin_B.
DR   InterPro; IPR002049; Laminin_EGF.
DR   Pfam; PF00052; laminin_B; 1.
DR   Pfam; PF00053; laminin_EGF; 10.
DR   Pfam; PF00055; laminin_Nterm; 1.
DR   ProDom; PD003031; Laminin_B.
DR   ProDom; PD002082; LamNT; 1.
DR   SMART; SM00180; EGF_Lam; 10.
DR   SMART; SM00001; EGF_like; 1.
DR   SMART; SM00281; LamNT; 1.
DR   SMART; SM00136; LamB; 1.
DR   PROSITE; PS00022; EGF_1; 8.
DR   PROSITE; PS01186; EGF_2; 11.
DR   PROSITE; PS01248; LAMININ_TYPE_EGF; 11.
KW   Glycoprotein; Basement membrane; Extracellular matrix; Coiled coil;
KW   Laminin EGF-like domain; Cell adhesion; Repeat; Signal.
FT   SIGNAL           1     33       POTENTIAL.
FT   CHAIN           34   1639       LAMININ GAMMA-1 CHAIN.
FT   DOMAIN          34    298       LAMININ N-TERMINAL (DOMAIN VI).
FT   DOMAIN         299    523       4.5 X LAMININ EGF-LIKE REPEATS (DOMAIN
FT                                    V).
FT   DOMAIN         299    358       LAMININ EGF-LIKE 1.
FT   DOMAIN         359    413       LAMININ EGF-LIKE 2.
FT   DOMAIN         414    460       LAMININ EGF-LIKE 3.
FT   DOMAIN         461    513       LAMININ EGF-LIKE 4.
FT   DOMAIN         514    523       LAMININ EGF-LIKE 5 (N-TERMINAL).
FT   DOMAIN         524    709       LAMININ DOMAIN IV.
FT   DOMAIN         710   1049       6.5 X LAMININ EGF-LIKE REPEATS (DOMAIN
FT                                    III).
FT   DOMAIN         710    743       LAMININ EGF-LIKE 5 (C-TERMINAL).
FT   DOMAIN         744    792       LAMININ EGF-LIKE 6.
FT   DOMAIN         793    846       LAMININ EGF-LIKE 7.
FT   DOMAIN         847    901       LAMININ EGF-LIKE 8.
FT   DOMAIN         902    955       LAMININ EGF-LIKE 9.
FT   DOMAIN         956   1003       LAMININ EGF-LIKE 10.
FT   DOMAIN        1004   1049       LAMININ EGF-LIKE 11.
FT   DOMAIN        1050   1087       DOMAIN II AND I.
FT   DOMAIN        1109   1247       DOMAIN II AND I.
FT   DOMAIN        1306   1627       COILED COIL (POTENTIAL).
FT   DISULFID       299    308       BY SIMILARITY.
FT   DISULFID       301    322       BY SIMILARITY.
FT   DISULFID       324    333       BY SIMILARITY.
FT   DISULFID       336    356       BY SIMILARITY.
FT   DISULFID       359    368       BY SIMILARITY.
FT   DISULFID       361    384       BY SIMILARITY.
FT   DISULFID       387    396       BY SIMILARITY.
FT   DISULFID       399    411       BY SIMILARITY.
FT   DISULFID       414    426       BY SIMILARITY.
FT   DISULFID       416    432       BY SIMILARITY.
FT   DISULFID       434    443       BY SIMILARITY.
FT   DISULFID       446    458       BY SIMILARITY.
FT   DISULFID       461    475       BY SIMILARITY.
FT   DISULFID       463    482       BY SIMILARITY.
FT   DISULFID       484    493       BY SIMILARITY.
FT   DISULFID       496    511       BY SIMILARITY.
FT   DISULFID       744    753       BY SIMILARITY.
FT   DISULFID       746    760       BY SIMILARITY.
FT   DISULFID       762    771       BY SIMILARITY.
FT   DISULFID       774    790       BY SIMILARITY.
FT   DISULFID       793    801       BY SIMILARITY.
FT   DISULFID       795    811       BY SIMILARITY.
FT   DISULFID       814    823       BY SIMILARITY.
FT   DISULFID       826    844       BY SIMILARITY.
FT   DISULFID       847    861       BY SIMILARITY.
FT   DISULFID       849    868       BY SIMILARITY.
FT   DISULFID       871    880       BY SIMILARITY.
FT   DISULFID       883    899       BY SIMILARITY.
FT   DISULFID       902    919       BY SIMILARITY.
FT   DISULFID       904    926       BY SIMILARITY.
FT   DISULFID       928    937       BY SIMILARITY.
FT   DISULFID       940    953       BY SIMILARITY.
FT   DISULFID       956    968       BY SIMILARITY.
FT   DISULFID       958    975       BY SIMILARITY.
FT   DISULFID       977    986       BY SIMILARITY.
FT   DISULFID       989   1001       BY SIMILARITY.
FT   DISULFID      1004   1016       BY SIMILARITY.
FT   DISULFID      1006   1022       BY SIMILARITY.
FT   DISULFID      1024   1033       BY SIMILARITY.
FT   DISULFID      1036   1047       BY SIMILARITY.
FT   DISULFID      1050   1053       INTERCHAIN (PROBABLE).
FT   DISULFID      1050   1053       INTERCHAIN (PROBABLE).
FT   DISULFID      1631   1631       INTERCHAIN (PROBABLE).
FT   CARBOHYD       147    147       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD      1631   1631       N-LINKED (GLCNAC...) (POTENTIAL).
```

```
Query Match             8.6%;  Score  159.5;  DB  1;  Length 1639;
Best Local Similarity  23.9%;  Pred. No. 2;
Matches  105;  Conservative  58;  Mismatches  162;  Indels  115;  Gaps  19;

Qy      2 ACNNSGKDGNTSANSAD-------ESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDE 55
            : :||| |||| ||  ||        ||| :||| ||  |  ||:  ||  |: |:  | ||
Db   1107 ARDNSGDGGQTYAEVIDDLHKHLDSVR-EHLVSADKFQADANGEIDRARQNYTILDQITE 1165

Qy     56 IAAKAIGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCS 115
           | | |  ||  | |    |:   |  ||  | | |: |:| | :|| | ||||
Db   1166 NAKKEL-----QQALDLLNDEGAQALARAKEKSVEFCQQ------------SEQISDISREA 1210

Qy    116 ETFTNKLK-EKHTDLGKEGVTDADAKEAILKAN----------GTKTKG-AEELGK 159
           ||:| |:|  :  |: :|    :  | | |  |            |: | |  || | |
Db   1211 RALADKLESEAQFDLKNA-----KDAKDAVEKAHQLAKSAIDLQLKIGTELRSEVGLELSH 1266
```

```
Qy   160  LFESVEVLSKAAKE------------------MLANSVKELTSPVVAESPKKPSMVNNSGKDGNT 206
               ||:|:|:||                     ||
Db  1267  VKOSLGTVVQTSKEALRKANEVYDTALTLLNDVNRQTQPEIDISQLK------------KDAVA 1318

Qy   207  SANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDELATKAIGKKIQQ------ 261
               |||:|                                                  |
Db  1319  ANERADELLK--QITELS---NSNGELFADFEFEQEL----TEALLKRAEOQOLED 1365

Qy   262  -----------NGGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSE---------- KL 299
                                                                       :|
Db  1366  IELLERAKAAHDKATKAVEQGDN-TLKEANNTYEKLAGPQSDVQRSSESAEKALQTVPNI 1424

Qy   300  KEKIENAKKCSEDFTKKLEGEHAQLGIENVTDENAKKAIL--ITDAAKDKGAAELEKLFK 357
          ::::||:|  |||||||||||||||                  ||  |             ||
Db  1425  EKEIQNAESLISQAEEEALDGANKN---ANEAAKNAQEAQLKYAEQASKD-----AELIR 1475

Qy   358  AVENLAKAAAKEMLANSVKEL 377
               | || :|
Db  1476  RKANETKVAARNLREEADQL 1495
```

RESULT 11

```
ID   YDB6_SCHPO     STANDARD;      PRT;   1957 AA.
AC   Q10411;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   HYPOTHETICAL 222.8 KDA PROTEIN C1F3.06C IN CHROMOSOME I.
GN   SPAC1F3.06C.
OS   Schizosaccharomyces pombe (fission yeast).
OC   Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC   Schizosaccharomycetales; Schizosaccharomycetaceae;
OC   Schizosaccharomyces.
OX   NCBI_TaxID=4896;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=972;
RA   Connor R., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;
RL   Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.
```

```
DR   EMBL; Z70690; CAA94624.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE 1957 AA; 222785 MW; 3F480CA0617D9DA CRC64;
```

```
Query Match           8.6%;  Score 159.5;  DB 1;  Length 1957;
Best Local Similarity 22.7%;  Pred. No. 2.4;
Matches 107;  Conservative 73;  Mismatches 151;  Indels 141;  Gaps 21;
```

```
Qy   8    KDGNTSANSADE---SVKGPNLTEINKKITDSNAVLLAV-KEVEALLSSIDEI------ 56
          |||  ||:||:|| :  :    || ||:| ||||| :|||  ::| ||| |
Db   360  KDSRTSNSQLEEAVELKESNRT-IHSQLTDAESKLSSFEQENKSLKGSIDEYQNNLSSK 418

Qy   57   ------------AAKAIGKKIHQNNGLDTEN------------ NHNGSLL 82
                       |:| | | ::  :  ||:||
Db   419  DKMVKQVSSQLEEARSSLAHATGKLAEINSERDFQNKKIKDFEKIEQDLRACLNSSSNEL 478

Qy   83   AGAVAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNKL------- KEKH 126
               | |:|  : :|||:  |  | | | |:|     |:||           |||
Db   479  KEK---SALIDKKDQELNN--LREQIKEQKKVSESTOSSLQSLQRDILNEKKHHEVYESQ 533

Qy   127  ----------TDLGCKEGYTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAA 171
```

```
Db   534  LNELKGELQTEISNSEHLSSQLSTLAAEKEAAVATN-------NELSESKNSLQTLCNAF 586
               ::|   | :|     | :: ||| |  : :              :|::|| | |
Qy   172  KEMLANSVKELTSPVVAESPKKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESN 231

Db   587  QEKLAKSVMQL----------------KENEQNFSSLDTSFK--KLNESHQELENNH 625
          :|:|| |                      |  || :|  :|   :||:: :|| :::
Qy   232  AVVLAVKEIETLLASIDELATKAIGKKIQQNGGLAVEAGHNGTLLAGAYTISKLITQK-- 289

Db   626  QTI--TKQLKDTSSKLQQLQLERANEE-QKESTLSDENNDLRTKLLKLEESNKSLIRKQE 682
          |:|    |:|     :  :::         |::|:::: ::|  ||:|: :|||:| ||
Qy   290  -LDGL-KNSEKLKE--KIENAKKCSE------DFTKKLEGEHAQLGIE-----NVTD 331

Db   683  DVDSLEKNIQTLKEDLRKSEEALRFSKLEAKNLREVIDNLKGKHETLEAQRNDLHSSLSD 742
          : |:|   :| ||:||  |: || |||           :||  |:  | |   | :: |
Qy   332  ENAKKAILITDAAKDKGAAELEKLFKRAVENLAKAAAKEM------LANSVKEL 377

Db   743  AKNTNAILSSELT--KSSEDVKRLTANVETLTQDSRAMKQSFTSLVNSYQSI 792
```

RESULT 12

```
ID   M6_STRPY       STANDARD;      PRT;    483 AA.
AC   P08089;
DT   01-AUG-1988 (Rel. 08, Created)
DT   01-AUG-1988 (Rel. 08, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   M PROTEIN, SEROTYPE 6 PRECURSOR.
GN   EMM6.
OS   Streptococcus pyogenes.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC   Streptococcus.
OX   NCBI_TaxID=1314;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=85166224; PubMed=3885219;
RA   Scott J.R., Pulliam W.M., Hollingshead S.K., Fischetti V.A.;
RT   "Relationship of M protein genes in group A streptococci.";
RL   Proc. Natl. Acad. Sci. U.S.A. 82:1822-1826(1985).
RN   [2]
RP   SEQUENCE OF 43-122 FROM N.A.
RX   MEDLINE=86111835; PubMed=3511046;
RA   Hollingshead S.K., Fischetti V.F., Scott J.R.;
RT   "Complete nucleotide sequence of type 6 M protein of the group A
RT   Streptococcus. Repetitive structure and membrane anchor.";
RL   J. Biol. Chem. 261:1677-1686(1986).
CC   -!- FUNCTION: THIS PROTEIN IS ONE OF THE DIFFERENT ANTIGENIC SEROTYPES
CC       OF M PROTEIN. M PROTEIN IS CLOSELY ASSOCIATED WITH VIRULENCE OF
CC       THE BACTERIUM AND CAN RENDER THE ORGANISM RESISTANT TO
CC       PHAGOCYTOSIS.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. CELL WALL.
CC   -!- SIMILARITY: TO OTHER M PROTEINS.
CC   -!- SIMILARITY: TO OTHER STREPTOCOCCAL AND STAPHYLOCOCCAL PROTEINS
CC       IN THE REGION OF THE MEMBRANE ANCHOR.
```

```
DR   EMBL; M11338; AAA26920.1; -.
DR   PIR; A26297; A26297.
DR   Interpro; IPR001899; Gram_pos_anchor.
DR   Interpro; IPR003345; M_repeat.
DR   Pfam; PF00746; Gram_pos_anchor; 1.
DR   Pfam; PF02370; M; 9.
DR   PRINTS; PR00015; GPOSANCHOR.
DR   PROSITE; PS00015; GPOSANCHOR.
DR   PROSITE; PS00343; GRAM_POS_ANCHORING; 1.
KW   Virulence; Phagocytosis; Cell wall; Duplication; Repeat; Antigen;
KW   Transmembrane; Coiled coil; Signal.
FT   SIGNAL        1     42
```

```
FT   CHAIN        43   483       M PROTEIN, SEROTYPE 6.
FT   DOMAIN       43   457       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    458   477       MEMBRANE ANCHOR (POTENTIAL).
FT   DOMAIN      478   483       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       69   138       10 X 7 AA TANDEM REPEATS.
FT   DOMAIN      157   269       4.5 X 25 AA TANDEM REPEATS.
FT   DOMAIN      279   347       TWO DIRECTLY REPEATED 27 AMINO ACID
FT   DOMAIN      348   411       BLOCKS SEPARATED BY 15 AMINO ACIDS.
FT   DOMAIN      412   448       HYDROPHILIC.
FT   DOMAIN      449   454       GLY/PRO-RICH (CELL WALL-SPANNING).
CC                                CONSERVED IN GRAM-POSITIVE COCCI SURFACE
CC                                PROTEINS.
SQ   SEQUENCE   483 AA;  53472 MW;  68F87F28DB53A448 CRC64;

Query Match             8.5%;  Score 157;  DB 1;  Length 483;
Best Local Similarity  23.0%;  Pred No. 0.67;
Matches   92;  Conservative   64;  Mismatches  144;  Indels  100;  Gaps  18;

Qy    5 NSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSSIDEIAAKAIGKK 64
Db   73 NNDKLTTENNNLTDON---KNITENKNLTDQNKNLTTENK------------------NL 112

Qy   65 IHQNNGLDPTENNINGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNKLKE 124
Db  113 TDQNKNLTTENK---ELKAEENRLTI------ENKGLTKKLSEAE--EEAANKERE 157

Qy  125 KHTDLGK-EGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKELT 183
Db  158 NKEAIGTLKKTLDETVKDKIAKEOESK----ETIGTL 200

Qy  184 SPVVAESPKKPKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEI--E 241
Db  201 -KIAKEOESKEII-------GTLKKTLDETVK----DKIAKE-OESKETIGTLKKILDE 246

Qy  242 TLLASI----------DELATKAIGKKIQONGGLAVEAGHNGTLLAGAYTISKL-- 285
Db  247 TVKDKIAREOKSKQDIGALKQELAKKDEGNKVSEASRKGLIRRDLDASREAKKQVEKDLAN 306

Qy  286 ITQKLDGLKNSEK--------LKEKIENAKKCSEDFTKKLEGEHAQLGIENVTDENAKKA 337
Db  307 LTAELDKVKEEKQISDASRGLIRRDLDASREAKKQVEKALEEANSKLAALEKLNKELEES 366

Qy  338 ILITDAAKDKGAAELEKLFKAVENLAKAAKEMLANSVKEL 377
Db  367 KKLTEKEK----AELQAKLEAE--AKALKEQLAKQAEEL 399
```

RESULT 13

```
ID   MYSN_DROME     STANDARD;      PRT;   2017 AA.
AC   Q99323;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   MYOSIN HEAVY CHAIN, NON-MUSCLE (ZIPPER PROTEIN) (MYOSIN II).
GN   ZIP.
OS   Drosophila melanogaster (fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RX   MEDLINE=90349606; PubMed=2117279;
RA   Ketchum A.S., Stewart C.T., Stewart M., Kiehart D.P.;
RT   "Complete sequence of the Drosophila nonmuscle myosin heavy-chain
RT   transcript: conserved sequences in the myosin tail and differential
RT   splicing in the 5' untranslated sequence.";
RL   Proc. Natl. Acad. Sci. U.S.A. 87:6316-6320(1990).
-!- FUNCTION: NONMUSCLE MYOSIN APPEARS TO BE RESPONSIBLE FOR
    MORPHOGENESIS AND CYTOKINESIS.
    CELLULARIZATION. REQUIRED FOR
```

```
-!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
    SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
-!- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
-!- SIMILARITY: CONTAINS 1 IQ DOMAIN.
```

```
DR   EMBL; M35012; AAA28713.1; -.
DR   PIR; A36014; A36014.
DR   PIR; B36014; B36014.
DR   HSSP; P08799; 1MND.
DR   FlyBase; FBgn0005634; zip.
DR   InterPro; IPR000048; IQ.
DR   InterPro; IPR002928; Myosin_tail.
DR   InterPro; IPR001609; myosin_head.
DR   Pfam; PF00612; IQ; 1.
DR   Pfam; PF00063; myosin_head; 1.
DR   Pfam; PF01576; Myosin_tail; 1.
DR   PRINTS; PR00193; MYOSINHEAVY.
DR   ProDom; PD000355; myosin_head; 1.
DR   SMART; SM00015; IQ; 1.
DR   SMART; SM00242; MYSc; 1.
DR   PROSITE; PS50096; IQ; 1.
KW   Myosin; Alternative splicing; Coiled coil; Actin-binding;
KW   ATP-binding; Calmodulin-binding.
FT   DOMAIN        1   829       MYOSIN HEAD-LIKE.
FT   DOMAIN      830   859       IQ.
FT   DOMAIN      886  2017       COILED COIL (POTENTIAL).
FT   NP_BIND     225   232       ATP.
FT   DOMAIN      250   260       25 KDA/50 KDA JUNCTION.
FT   DOMAIN      682   694       50 KDA/20 KDA JUNCTION.
FT   DOMAIN      705   727       ACTIN-BINDING.
FT   DOMAIN      742   758       REACTIVE SULFHYDRYL/ACTIN-BINDING.
FT   DOMAIN     1303  2017       LIGHT MEROMYOSIN (LMM).
FT   DOMAIN     1303  2017       ALPHA-HELICAL TAILPIECE.
FT   DOMAIN     1970  2017       GLOBULAR TAILPIECE.
FT   DOMAIN     1971  2017       GLOBULAR TAILPIECE (LMM).
FT   VARSPLIC      1    45       MISSING (IN SHORT ISOFORM).
SQ   SEQUENCE   2017 AA;  232016 MW;  73E3CB02BA8F2528 CRC64;

Query Match             8.2%;  Score 151.5;  DB 1;  Length 2017;
Best Local Similarity  20.6%;  Pred No. 5.7;
Matches   92;  Conservative   74;  Mismatches  158;  Indels  123;  Gaps  19;

Qy   10 GNTSANSADESVKGPNLTEIN--------------KKITDSNAVLLAVK----EVEALLSSSIDEIAA 58
Db 1271 GTLEAENADLATE---LRSVNSSRQENDRRKQAESQIAELQVKLAEIERARSELQEKCT 1327

Qy   59 KAIGKKIHQNNGLD---------------TENNHNGSLLAGAYAISTLIKQKLDGL------- 99
Db 1328 KLQQEAENITNQLEEAELKASAAVKSASNMESQLTEAQQLLEEETRQKL-GLSSKLRQIE 1386

Qy  100 -KNEGLKEKIDAAKKCSETFTNKLEKHT---DIGKEGVTDAD-AKEAILKANGTKTKGA 154
Db 1387 SEKEALQEDLEEDDEAKRNYERKLAEVTTQMQEIKKAEEDADLAKEL------EEGK 1438

Qy  155 EELGKLFESVEVLSKAAKEMLANSVK-----------ELISPVVAESPKKPKPSMVNNSGKDG 204
Db 1439 KRLNK---DIEALEROVKELIAQNDRLDKSKKKIQSELEDATIELEAQRTKVLELEKKQK 1495

Qy  205 NTSANSADESVKGPNLTE----ISKKITESNAVVLAV-KEIETLLASIDELATKAIGKKI 259
Db 1496 NFDKLIAEEKAISEQIAQERDTAEREAREKETKVLSVSRELEDEAFDKIEDLENK---RKT 1552

Qy  260 QONGGLAVEAGHNGTLLAGAYTISKLITOKLDGLKNSERLKEK----IENAKKCSEDFTK 315
Db 1553 LQN--------------------------ELDDLANTQGTADKNVHELEKAKRALESQLA 1586
```

```
Qy   316 KLEGEHAQLGIENVTDENAKKAILITD------------------AAKDKGAAELEK-LFKAVENL- 362
              ::::: |||     |||   |: |   ::||:||    : :|:: |  ||:   :| :|
Db  1587 ELKAQNEELEDDLQLTEDAKLRLEVNMQALRSQFERDLLAKEEGAEERRGLVKQLRDLE 1646

Qy   363 ----------AKAAKEMLANSVKEL 377
                    :| : |:::  :||
Db  1647 TELDEERKQRTAAVASKKKLEGDLKEI 1673
```

RESULT 14
ID   MYS_AEQIR   STANDARD;   PRT;   1938 AA.
AC   P24733;
DT   01-MAR-1992 (Rel. 21, Created)
DT   01-MAR-1992 (Rel. 21, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   MYOSIN HEAVY CHAIN, STRIATED MUSCLE.
OS   Aequipecten irradians (Bay scallop).
OC   Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorphia; Pectinoida;
OC   Pectinoidea; Pectinidae; Argopecten.
OX   NCBI_TaxID=31199;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Adductor muscle;
RX   MEDLINE=92011595; PubMed=1917970;
RA   Nyitray L., Goodwin E.B., Szent-Gyoergyi A.G.;
RT   "Complete primary structure of a scallop striated muscle myosin heavy
RT   chain. Sequence comparison with other heavy chains reveals regions
RT   that might be critical for regulation.";
RL   J. Biol. Chem. 266:18469-18476(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Adductor muscle;
RX   MEDLINE=91088319; PubMed=2263488;
RA   Nyitray L., Goodwin E.B., Szent-Gyoergyi A.G.;
RT   "Nucleotide sequence of full length cDNA for a scallop striated
RT   muscle myosin heavy chain.";
RL   Nucleic Acids Res. 18:7158-7158(1990).
RN   [3]
RP   X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 777-836.
RX   MEDLINE=94173332; PubMed=8127365;
RA   Xie X., Harrison D.H., Schlichting I., Sweet R.M., Kalabokis V.N.,
RA   Szent-Gyorgyi A.G., Cohen C.;
RT   "Structure of the regulatory domain of scallop myosin at 2.8-A
RT   resolution.";
RL   Nature 368:306-312(1994).
RN   [4]
RP   X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 777-836.
RX   MEDLINE=96419133; PubMed=8805510;
RA   Houdusse A., Cohen C.;
RT   "Structure of the regulatory domain of scallop myosin at 2-A
RT   resolution: Implications for regulation.";
RL   Structure 4:21-32(1996).
CC   -!- FUNCTION: MUSCLE CONTRACTION.
CC   -!- FUNCTION: MYOSIN IS A PROTEIN THAT BINDS TO F-ACTIN & HAS ATPASE
CC   ACTIVITY THAT IS ACTIVATED BY F-ACTIN.
CC   -!- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
CC   HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
CC   AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
CC   -!- SUBCELLULAR LOCATION: THICK FILAMENTS OF THE MYOFIBRILS.
CC   -!- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 IQ DOMAIN.
```

```
DR   EMBL; X55714; CAA39247.1; -.
DR   PIR; S13557; S13557.
DR   PIR; A40997; A40997.
DR   PDB; 1SCM; 30-APR-94.
DR   PDB; 1WDC; 11-JUL-96.
DR   SMART; SM00242; MYSc; 1.
DR   SMART; SM00015; IQ; 1.
DR   PROSITE; PS50096; IQ; 1.
DR   InterPro; IPR000048; IQ.
DR   InterPro; IPR002928; Myosin_tail.
DR   InterPro; IPR020017; Spectrin.
DR   InterPro; IPR001609; myosin_head.
DR   Pfam; PF00063; myosin_head; 1.
DR   Pfam; PF01576; Myosin_tail; 1.
DR   Pfam; PF00612; IQ; 1.
DR   PRINTS; PR00193; MYOSINHEAVY.
DR   ProDom; PD0000355; myosin_head; 1.
KW   Muscle protein; Coiled coil; Thick filament; Actin-binding;
KW   ATP-binding; Alkylation; Calmodulin-binding; 3D-structure.
FT   DOMAIN        1    777    MYOSIN HEAD-LIKE.
FT   DOMAIN      778    805    IQ.
FT   DOMAIN      836   1938    RODLIKE TAIL (S2 AND LMM DOMAINS).
FT   DOMAIN      836   1938    COILED COIL (POTENTIAL).
FT   NP_BIND     176    183    ATP (BY SIMILARITY).
FT   MOD_RES     693    693    ALKYLATION (SH-1) (BY SIMILARITY).
FT   MOD_RES     703    703    ALKYLATION (SH-2) (BY SIMILARITY).
FT   HELIX       778    821
FT   TURN        822    823
FT   HELIX       825    833
FT   HELIX       834    835
FT   TURN        834    835
SQ   SEQUENCE   1938 AA;  222821 MW;  A5CCE4127D1A4896 CRC64;
```

```
Query Match             8.1%;
Best Local Similarity   20.4%;
Matches    81;  Conservative   65;  Mismatches  166;  Indels   86;  Gaps   12;
                        Score 150;  DB 1;  Length 1938;
                        Pred. No. 6.4;
```

```
Qy   18 DESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGKKIHQNNGLDT----- 73
Db  925 DEEDAAADLEGTKKKMEADNA-----NLKDIGDLENTLQKAEQDKAHKDNQISTLQGE 978

Qy   74 ---ENNHNGSLLAGAYAISTLIKQKLDGLIKNEGLKEKIDAAKKCSEFTNKILKEKHTDLG 130
Db  979 ISQODEHIGKLNEKKALEEANKKTSDSLQAE--EDKCNHLNKLKAKLEQALDELEDNLE 1036

Qy  131 KEGVTDADAKEAILKANGTKITKGAEELGKLFESVEVLSKAAKEMLANSVKELTSPVVAES 190
Db 1037 REKKVRGDVERA-------KRRVEQDLKSTQENVEDLERVKRELEEN-------VRRK 1080

Qy  191 PKKPSMVNNSGKDGNTSANSADESVK--GPNLTEISKKI-TESNAVVLAVKEIETLLASI 247
Db 1081 EAEISSLNSKLEDEQNLVSQLQRKIKELQARIEELEELEAERNARAKVEKQRAELNREL 1140

Qy  248 DELATKRAIGKKIQQNGG--------------------LAVEAGHNGTLLAGAYTISKL----- 285
Db 1141 EEL-----GERLDEAGGATSAQIELNKKREAELLKIRRDLEEASLQHEAQISALRKKHQD 1195

Qy  286 -ITQKILDGLKNSEKLKEKIENAKK-----------------------CSEDFTKKLEGEHAQ 323
Db 1196 AANEMADQVDLQKVKSKLEKDKKDLKREMDDLESQMTHNMKNKGCSEKVMKQFESQMSD 1255

Qy  324 LGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVEN 361
Db 1256 L---NARLEDSQRSINELQSQKSRLQAENSDITRQLED 1290
```

RESULT 15
ID   MYSB_CAEEL   STANDARD;   PRT;   1966 AA.
AC   P02566;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   MYOSIN HEAVY CHAIN B (MHC B).

```
GN   UNC-54 OR MYO-4.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=82273600; PubMed=6576334;
RA   Karn J., Brenner S., Barnett L.;
RT   "Protein structural domains in the Caenorhabditis elegans unc-54
RT   myosin heavy chain gene are not separated by introns.";
RL   Proc. Natl. Acad. Sci. U.S.A. 80:4253-4257(1983).
RN   [2]
RP   SEQUENCE OF 850-1966 FROM N.A.
RX   MEDLINE=82272395; PubMed=7202124;
RA   McLachlan A.D., Karn J.;
RT   "Periodic charge distributions in the myosin rod amino acid sequence
RT   match cross-bridge spacings in muscle.";
RL   Nature 299:226-231(1982).
RN   [3]
RP   SEQUENCE OF 1876-1966 FROM N.A.
RX   MEDLINE=83238692; PubMed=6571695;
RA   Wills N., Gesteland R.F., Karn J., Barnett L., Bolten S.,
RA   Waterston R.H.;
RT   "The genes sup-7 X and sup-5 III of C. elegans suppress amber
RT   nonsense mutations via altered transfer RNA.";
RL   Cell 33:575-583(1983).
CC   -!- FUNCTION: MUSCLE CONTRACTION.
CC   -!- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
CC       HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
CC       AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
CC   -!- SUBCELLULAR LOCATION: THICK FILAMENTS OF THE MYOFIBRILS.
CC   -!- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CC       CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CC       CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC   -!- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
CC       ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
CC   -!- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
CC       MEROMYOSIN (LMM) AND 1 HEAVY MEROMYOSIN (HMM). IT CAN LATER BE
CC       SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
CC       SUBFRAGMENT (S2).
CC   -!- MISCELLANEOUS: THERE ARE FOUR DIFFERENT MYOSIN HEAVY CHAINS IN C.
CC       ELEGANS.
CC   -!- MISCELLANEOUS: MHC A AND MHC B ARE FOUND EXCLUSIVELY IN THE BODY
CC       WALL MUSCLE. THEY CO-ASSEMBLE INTO BODY WALL THICK FILAMENT.
CC   -!- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; J01050; AAA28124.1; -.
DR   EMBL; V01494; CAA24738.1; -.
DR   PIR; A02992; MWKW.
DR   HSSP; P08799; 1MND.
DR   InterPro; IPR002928; Myosin_tail.
DR   InterPro; IPR001609; myosin_head.
DR   Pfam; PF00063; myosin_head; 1.
DR   Pfam; PF01576; myosin_tail; 1.
DR   PRINTS; PR00193; MYOSINHEAVY.
DR   ProDom; PD000355; myosin_head; 1.
DR   SMART; SM00242; MYSc; 1.
KW   Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
KW   ATP-binding; Methylation; Alkylation; Multigene family.
FT   DOMAIN         .     850    MYOSIN HEAD-LIKE.
FT   DOMAIN        851    1966    COILED COIL (POTENTIAL).
FT   DOMAIN        851    1164    ALPHA-HELICAL TAILPIECE (S2).
FT   DOMAIN       1165    1176    HINGE.
FT   DOMAIN       1165    1966    LIGHT MEROMYOSIN (LMM).
```

```
FT   NP_BIND      177    184    ATP (BY SIMILARITY).
FT   DOMAIN       665    687    ACTIN-BINDING.
FT   DOMAIN       769    783    ACTIN-BINDING.
FT   DOMAIN       128    128    ACTIN-BINDING.
FT   MOD_RES      705    705    METHYLATION (TRI-) (POTENTIAL).
FT   MOD_RES      715    715    ALKYLATION (SH-1).
FT   MOD_RES     1337   1337    ALKYLATION (SH-2).
FT   CONFLICT    1880   1880    E -> R (IN REF. 2).
FT   CONFLICT    1880   1880    I -> L (IN REF. 2).
SQ   SEQUENCE    1966 AA;  225125 MW;  B66F0BB2FE27B67F CRC64;
```

```
Query Match             8.1%;  Score 150;  DB 1;  Length 1966;
Best Local Similarity  22.7%;  Pred. No. 6.5;
Matches   90;  Conservative  76;  Mismatches 149;  Indels  82;  Gaps  20;

Qy     16  SADESVKGPNLITELNKKIIFDSNAVLLAVK---EVEALLSSIDEI--AAKAIGKKIHQ-NN    69
Db   1358  SLEEEIEGKN--EILRQLSKANADIQQWKARFEGEGLLKA-DELEDAKRRQAQKINELQE    1414

Qy     70  GLDTENNINGSL-LAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFINKLKEKHTD    128
Db   1415  ALDAANSKNASLEKTKSRLVGDLDDAQVDVERANGVASALEKKKQKGFDKIIDEWRKKTDD   1474

Qy    129  LGKEGVTDADAKEAILKANGT---KTKGAEELGKLFESVEVLSKAAKEMLANSVKELTSP    185
Db   1475  LAAE----LDGAQRDLRNTSTDLFKARNAQE--ELAEVVEGLRRENKS-LSQEIKDLTDQ   1527

Qy    186  VVAESPKKPSMVNNSGKDGNTSANSADESVKGPNL-TEISKKITESNAVVLAVKEIETL    243
Db   1528  L-----------------GEGGRSVHEMQKIIRRLEIEKEELQHALDEAEA-ALEAEESKVL 1571

Qy    244  LASIDELATKA-IGKKIQQ----------------NGGLAVEAGHNGTLLAGAYTIS     283
Db   1572  RAQVEVSQIRSEIEKRIQEKEEEFENTRKNHARALESMQASLETEAKGKAELLR-----   1625

Qy    284  KLITQKLDGLKRNSEKLKEKIENAKKCSEDFTKKLEGEHAQLGIENVTDENAKKAILITDA  343
Db   1626  --IKKLEG--DINELEIALDHANKANADAQQKNLKR------------YQEQVRELQLQVEE 1671

Qy    344  AKDKGAAAELEKLFKAVE--NLAKAAKE--MLANSVKE    376
Db   1672  EQRNGADTREQFFNAEKRATLLQSEKEELLVANEAAE   1708
```

```
Search completed: March 18, 2002, 10:11:49
Job time: 967 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 10:10:53 ; Search time 124.19 Seconds
(without alignments)
444.035 Million cell updates/sec

Title: US-09-596-746A-30
Perfect score: 1850
Sequence: 1 MACNNSGKDGNTSANSADES.........AVENLAKAAKEMLANSVKEL 377

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :   SPTREMBL_17:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	943	51.0	192	2	Q9S3P3	Q9s3p3 borrelia bu
2	927	50.1	191	2	P70818	P70818 borrelia bu
3	921	49.8	192	2	Q9R7B1	Q9r7b1 borrelia bu
4	908	49.1	193	2	Q9RR53	Q9rr53 borrelia bu
5	893	48.3	200	2	Q9RQR9	Q9rqr9 borrelia bu
6	885	47.8	182	2	Q9R7B2	Q9r7b2 borrelia bu
7	876	47.4	193	2	P94234	P94234 borrelia bu
8	876	47.4	194	2	O31122	O31122 borrelia bu
9	876	47.4	201	2	Q9RQR8	Q9rqr8 borrelia bu
10	863	46.6	191	2	O44726	O44726 borrelia bu
11	853	46.1	177	2	Q9REH7	Q9reh7 borrelia bu
12	843.5	45.6	178	2	O31123	O31123 borrelia bu
13	841	45.5	185	2	O31123	O31123 borrelia bu
14	829	44.8	175	2	Q9R7B4	Q9r7b4 borrelia bu
15	826	44.6	178	2	O44995	O44995 borrelia bu
16	823	44.5	178	2	O44979	O44979 borrelia bu
17	802	43.4	163	2	Q9R7A9	Q9r7a9 borrelia bu
18	774.5	41.9	211	2	O44720	O44720 borrelia bu
19	773.5	41.8	193	2	P94237	P94237 borrelia bu
20	773	41.8	159	2	Q9R7B0	Q9r7b0 borrelia bu
21	739	39.9	178	2	O44994	O44994 borrelia bu
22	709	38.3	210	2	O44719	O44719 borrelia af
23	707	38.2	192	2	Q9S3P2	Q9s3p2 borrelia bu
24	696.5	37.6	211	2	O44977	O44977 borrelia bu
25	692.5	37.4	211	2	O49576	O49576 borrelia bu
26	690.5	37.3	193	2	P94242	P94242 borrelia bu
27	686.5	37.1	211	2	O49576	O49576 borrelia bu
28	686	37.1	190	2	P94244	P94244 borrelia bu
29	686	37.1	212	2	Q9K1M5	Q9k1m5 borrelia af
30	685.5	37.1	209	2	O44671	O44671 borrelia af
31	683.5	36.9	191	2	P94245	P94245 borrelia bu
32	683	36.9	212	2	O44705	O44705 borrelia bu
33	681	36.8	212	2	O31115	O31115 borrelia bu
34	675	36.5	194	2	P94247	P94247 borrelia bu
35	670	36.2	212	2	O44727	O44727 borrelia bu
36	667	36.1	201	2	P96571	P96571 borrelia ja
37	666.5	36.0	193	2	P94231	P94231 borrelia bu
38	666.5	36.0	211	2	O92C7	O92c7 borrelia va
39	666	36.0	202	2	P96573	P96573 borrelia ja
40	666	36.0	193	2	P94233	P94233 borrelia bu
41	665	35.9	202	2	P96571	P96571 borrelia ja
42	664.5	35.9	209	2	O92C7	O92c7 borrelia va
43	664.5	35.9	211	2	O44883	O44883 borrelia bu
44	664.5	35.9	203	2	O49577	O49577 borrelia af
45	664	35.9	212	2	O50624	O50624 borrelia af
					O44670	O44670 borrelia af

ALIGNMENTS

RESULT 1
Q9S3P3

```
ID   Q9S3P3      PRELIMINARY;      PRT;   192 AA.
AC   Q9S3P3;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C (FRAGMENT).
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=IP2;
RX   MEDLINE=96296448; PubMed=8709845;
RA   Livey I., Gibbs C.P., Schuster R., Dorner F.;
RT   "Evidence for lateral transfer and recombination in OspC variation in
RT   Lyme disease Borrelia.";
RL   Mol. Microbiol. 18:257-269(1995).
DR   EMBL; L42887; AAB36995.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
FT   NON_TER         1      1
FT   NON_TER       192    192
SQ   SEQUENCE   192 AA;  20287 MW;  11846F7AC84C7E3D CRC64;
```

Query Match   51.0%;   Score 943;   DB 2;   Length 192;
Best Local Similarity   99.0%;
Matches   190;   Conservative   1;   Mismatches   1;   Indels   0;   Gaps   0;

```
Qy    3 CNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAARAIG 62
          ||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db    1 CNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAARAIG 60

Qy   63 KKIHQNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNKL 122
          ||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db   61 KKIHQNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAARKCSETFTNKL 120

Qy  123 KEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSRAAKEMLANSVKEL 182
```

```
                      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   121 KEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKEL 180

Qy   183 TSPVVAESPKKP 194
                  ||||||||||||
Db   181 TSPVVAESPKKP 192
```

RESULT 2

```
ID   P70818       PRELIMINARY;      PRT;   191 AA.
AC   P70818;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN (FRAGMENT).
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=2-1498 CA4;
RA   Probert W.S., Crawford M.R., Cadiz R.B., Lefebvre R.B.;
RL   Submitted (SEP-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; L81131; AAB06569.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
FT   NON_TER        1      1
SQ   SEQUENCE   191 AA;  20126 MW;  D2B9B1CB2B4DC3C0 CRC64;
```

```
Query Match             50.1%;  Score 927;  DB 2;  Length 191;
Best Local Similarity   98.4%;
  Matches  188;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy     4 NNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKALGK 63
           |||||||||||||||||||||||||||  ||||||||||||||||||||||||||||| ||
Db     1 NNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGK 60

Qy    64 KIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNKIK 123
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db    61 KIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNKLK 120

Qy   124 EKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKELT 183
           ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db   121 EKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKELT 180

Qy   184 SPVVAESPKKP 194
           |||||||||||
Db   181 SPVVAESPKKP 191
```

RESULT 3

```
ID   Q9R7B1       PRELIMINARY;      PRT;   192 AA.
AC   Q9R7B1;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C (FRAGMENT).
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TET5;
RX   MEDLINE=9478003; PubMed=9336916;
RA   Ras N.M., Postic D., Foretz M., Baranton G.;
RT   "Borrelia burgdorferi sensu stricto, a bacterial species 'made in the
RT   U.S.A.'?";
RL   Int. J. Syst. Bacteriol. 47:1112-1117(1997).
```

```
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TET5;
RA   Marti-Ras N., Postic D., Foretz M., Baranton G.;
RL   Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U91798; AAB81895.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
FT   NON_TER        1      1
SQ   SEQUENCE   192 AA;  20297 MW;  6770502A20AAF764 CRC64;
```

```
Query Match             49.8%;  Score 921;  DB 2;  Length 192;
Best Local Similarity   97.9%;
  Matches  186;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy     1 MACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKA 60
         : :||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||
Db     3 ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA 62

Qy    61 IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    63 IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN 122

Qy   121 KLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 180
         ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db   123 KLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 182

Qy   181 ELTSPVVAES 190
         ||||||||||
Db   183 ELTSPVVAES 192
```

RESULT 4

```
ID   Q9RR53       PRELIMINARY;      PRT;   193 AA.
AC   Q9RR53;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C (FRAGMENT).
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=OC1;
RX   MEDLINE=99091544; PubMed=9872945;
RA   Wang I.N., Dykhuizen D.E., Qiu W., Dunn J.J., Bosler E.M., Luft B.J.;
RT   "Genetic diversity of ospc in a local population of Borrelia
RT   burgdorferi sensu stricto.";
RL   Genetics 151:15-30(1999).
DR   EMBL; AF029860; AAB86543.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
FT   NON_TER        1      1
SQ   SEQUENCE   193 AA;  20502 MW;  5EFDB5AFF8986D1E CRC64;
```

```
Query Match             49.1%;  Score 908;  DB 2;  Length 193;
Best Local Similarity   97.9%;
  Matches  183;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy     1 MACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKA 60
         : :||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||
Db     7 ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA 66

Qy    61 IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN 120
```

```
Db   67  IGKKIHQNNGLDTENNHNGSLLAGAYAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN  126
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy  121  KLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK  180
         |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db  127  KLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK  186

Qy  181  ELTSPVV  187
         |||||||
Db  187  ELTSPVV  193
```

RESULT 5
ID   Q9ROR9   PRELIMINARY;   PRT;   200 AA.
AC   Q9ROR9;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C PRECURSOR (FRAGMENT).
GN   OSPC.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OG   Plasmid cp26.
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=2E7;
RX   MEDLINE=20002545; PubMed=10531219;
RA   Hofmeister E.K., Glass G.E., Childs J.E., Persing D.H.;
RT   "Population dynamics of a naturally occuring heterogeneous mixture of
RT   Borrelia burgdorferi clones.";
RL   Infect. Immun. 67:5709-5716(1999).
DR   EMBL; AF074464; AAD23911.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
KW   Plasmid.
FT   NON_TER   200   200
SQ   SEQUENCE   200 AA; 21307 MW; A22F158758EBEB6B CRC64;

Query Match   48.3%; Score 893; DB 2; Length 200;
Best Local Similarity 97.8%; Pred. No. 5.7e-35;
Matches 180; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

```
Qy   1  MACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKA  60
        :: ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db  17  ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA  76

Qy  61  IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN  120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  77  IGKKIHQNNGLDTENNHNGSLLAGAYAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN  136

Qy  121  KLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK  180
         |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db  137  KLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK  196

Qy  181  ELTS  184
         ||||
Db  197  ELTS  200
```

RESULT 6
ID   Q9R7B2   PRELIMINARY;   PRT;   182 AA.
AC   Q9R7B2;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C (FRAGMENT).
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TETS;
RX   MEDLINE=97478003; PubMed=9336916;
RA   Ras N.M., Postic D., Foretz M., Baranton G.;
RT   "Borrelia burgdorferi sensu stricto, a bacterial species 'made in the
RT   U.S.A.'?";
RL   Int. J. Syst. Bacteriol. 47:1112-1117(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TETS;
RA   Marti-Ras N., Postic D., Foretz M., Baranton G.;
RL   Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U91797; AAB81894.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; -.
DR   ProDom; PD001149; Lipoprotein_6; 1.
SQ   SEQUENCE   182 AA; 19202 MW; 422146F99AA57BF2 CRC64;

Query Match   47.8%; Score 885; DB 2; Length 182;
Best Local Similarity 98.9%; Pred. No. 1.2e-34;
Matches 180; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy   7  GKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGKKIH  66
        |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db   1  GKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGKKIH  60

Qy  67  QNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNKLKEKH  126
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  61  QNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNKLKEKH  120

Qy  127  TDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKELTSPV  186
         |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db  121  TDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKELTSPV  180

Qy  187  VA  188
         ||
Db  181  VA  182
```

RESULT 7
ID   P94234   PRELIMINARY;   PRT;   193 AA.
AC   P94234;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C (FRAGMENT).
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=28354;
RX   MEDLINE=96296448; PubMed=8709845;
RA   Livey I., Gibbs C.P., Schuster R., Dorner F.;
RT   "Evidence for lateral transfer and recombination in OspC variation in
RT   Lyme disease Borrelia.";
RL   Mol. Microbiol. 18:257-269(1995).
DR   EMBL; L42895; AAB37003.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
FT   NON_TER   1   1
SQ   SEQUENCE   193 AA; 20411 MW; 05B68720F061E2A0 CRC64;

```
RESULT 8

ID   O31122      PRELIMINARY;   PRT;   194 AA.
AC   O31122;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C (FRAGMENT).
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Wang I.-N., Dykhuizen D.E., Dunn J.J., Luft B.J.;
RL   Submitted (OCT-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF029871; AAB86554.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
NON_TER   1
NON_TER   194
SQ   SEQUENCE   194 AA;   20640 MW;   00A5E6E2D2CE0F7F CRC64;

Query Match             47.4%;  Score 876;  DB 2;  Length 193;
Best Local Similarity  100.0%;  Pred. No. 3.4e-34;
Matches  180;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy   198 NNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDELATKAIGK 257
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2 NNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDELATKAIGK 61

Qy   258 KIQQNGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKEKIENAKKCSEDFTKKL 317
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    62 KIQQNGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKEKIENAKKCSEDFTKKL 121

Qy   318 EGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLAKAAKEMLANSVKEL 377
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   122 EGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLAKAAKEMLANSVKEL 181
```

```
RESULT 9

ID   Q9RQR8      PRELIMINARY;   PRT;   201 AA.
AC   Q9RQR8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C PRECURSOR (FRAGMENT).
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Plasmid cp26.
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=386';
RX   MEDLINE=20002545; PubMed=10531219;
RA   Hofmeister E.K., Glass G.E., Childs J.E., Persing D.H.;
RT   "Population dynamics of a naturally occurring heterogeneous mixture of
RT   Borrelia burgdorferi clones.";
RL   Infect. Immun. 67:5709-5716(1999).
DR   EMBL; AF074465; AAD23912.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
KW   Plasmid.
NON_TER   201
SQ   SEQUENCE   201 AA;   21457 MW;   806F198295101B07 CRC64;

Query Match             47.4%;  Score 876;  DB 2;  Length 194;
Best Local Similarity  100.0%;  Pred. No. 3.4e-34;
Matches  180;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy   198 NNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDELATKAIGK 257
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    10 NNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDELATKAIGK 69

Qy   258 KIQQNGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKEKIENAKKCSEDFTKKL 317
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    70 KIQQNGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKEKIENAKKCSEDFTKKL 129

Qy   318 EGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLAKAAKEMLANSVKEL 377
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   130 EGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLAKAAKEMLANSVKEL 189
```

```
Query Match             47.4%;  Score 876;  DB 2;  Length 201;
Best Local Similarity  100.0%;  Pred. No. 3.5e-34;
Matches  180;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy   198 NNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDELATKAIGK 257
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    20 NNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDELATKAIGK 79

Qy   258 KIQQNGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKEKIENAKKCSEDFTKKL 317
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    80 KIQQNGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKEKIENAKKCSEDFTKKL 139

Qy   318 EGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLAKAAKEMLANSVKEL 377
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   140 EGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLAKAAKEMLANSVKEL 199
```

```
RESULT 10

ID   Q44726      PRELIMINARY;   PRT;   191 AA.
AC   Q44726;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C (FRAGMENT).
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=297;
RX   MEDLINE=94314437; PubMed=8039891;
RA   Fung B.P., McHugh G.L., Leong J.M., Steere A.C.;
RT   "Humoral immune response to outer surface protein C of Borrelia
RT   burgdorferi in Lyme disease: role of the immunoglobulin M response in
RT   the serodiagnosis of early infection.";
RL   Infect. Immun. 62:3213-3221(1994).
DR   EMBL; U08284; AAA21460.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
NON_TER   1
SQ   SEQUENCE   191 AA;   20150 MW;   C49A4030F0A28717 CRC64;

Query Match             46.6%;  Score 863;  DB 2;  Length 191;
Best Local Similarity   99.4%;  Pred. No. 1.3e-33;
Matches  178;  Conservative   0;  Mismatches   1;  Indels   0;  Gaps   0;

Qy   199 NSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDELATKAIGKK 258
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 NSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDELATKAIGKK 60

Qy   259 IQQNGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKEKIENAKKCSEDFTKKLE 318
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 IQQNGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKEKIENAKKCSEDFTKKLE 120
```

```
Qy   319 GEHAQLGIENVTDENAKKAILTDAAKDKGAAELEKLFKAVENLAKAAKEMLANSVKEL 377
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLAKAAKEMLANSVKEL 179
```

RESULT 11
ID   Q44999
AC   Q44999;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C (FRAGMENT).
OSPC.
GN   OSPC.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TXGW;
RX   MEDLINE=95286481; PubMed=7768799;
RA   Theisen M., Borre M., Mathiesen M.J., Mikkelsen B., Lebech A.M.,
RA   Hansen K.;
RT   "Evolution of the Borrelia burgdorferi outer surface protein OspC.";
RL   J. Bacteriol. 177:3036-3044(1995).
FT   NON_TER   1   1
FT   NON_TER   177   177
SQ   SEQUENCE   177 AA; 18800 MW; 15667IB9614E7A2D CRC64;

Query Match   98.3%;   Score 853; DB 2; Length 177;
Best Local Similarity   98.3%;
Matches 174; Conservative   1; Pred. No. 3.6e-33; Mismatches   2; Indels   0; Gaps   0;

```
Qy    11 NTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGKKIHQNNG 70
         ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db     1 NTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGKKIHQNNG 60

Qy    71 LDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNKLKEKHTDLG 130
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 LDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNKLKEKHTDLG 120

Qy   131 KEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKELTSPVV 187
         ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db   121 KRGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKELTSPVV 177
```

DR   EMBL; X73622; CAA52001.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
FT   NON_TER   1   1
FT   NON_TER   178   178
SQ   SEQUENCE   178 AA; 18894 MW; D619AE4C646F14EB CRC64;

Query Match   45.6%;   Score 843.5; DB 2; Length 178;
Best Local Similarity   97.8%;
Matches 174; Conservative   1; Pred. No. 1e-32; Mismatches   2; Indels   1; Gaps   1;

```
Qy    11 NTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGKKIHQNNG 70
         ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db     1 NTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGKKIHQNNG 60

Qy    71 LDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNKLKEKHTD-L 129
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db    61 LDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNKLKEKHTDSF 120

Qy   130 GKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKELTSPVV 187
         ||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db   121 GKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKELTSPVV 178
```

RESULT 12
ID   Q9REH7
AC   Q9REH7;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
OSPC.
DE   OUTER SURFACE PROTEIN C (FRAGMENT).
GN   OSPC.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=B31;
RX   MEDLINE=94075528; PubMed=8253951;
RA   Theisen M., Frederiksen B., Lebech A.M., Vuust J., Hansen K.;
RT   "Polymorphism in ospC gene of Borrelia burgdorferi and
RT   immunoreactivity of OspC protein: implications for taxonomy and for
RT   use of OspC protein as a diagnostic antigen.";
RL   J. Clin. Microbiol. 31:2570-2576(1993).

PRELIMINARY;   PRT;   178 AA.

RESULT 13
ID   O31123
AC   O31123;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
OSPC.
DE   OUTER SURFACE PROTEIN C (FRAGMENT).
GN   OSPC.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=OC13;
RA   Wang I.-N., Dykhuizen D.E., Dunn J.J., Luft B.J.;
RL   Submitted (OCT-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF029872; AAB86555.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
FT   NON_TER   1   1
FT   NON_TER   185   185
SQ   SEQUENCE   185 AA; 19673 MW; 58D6FEE3C7769CAF CRC64;

Query Match   45.5%;   Score 841; DB 2; Length 185;
Best Local Similarity   97.7%;
Matches 172; Conservative   2; Pred. No. 1.4e-32; Mismatches   2; Indels   0; Gaps   0;

```
Qy   198 NNSGKDGNTSANSADESVKGPNLTEISKKITESNAVLLAVKEIETLLASIDELATKAIGK 257
         |||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db    10 NNSGKDGNTSANSADESVKGPNLTEISKKITESNAVLLAVKEIETLLASVDELATKAIGK 69

Qy   258 KIQQNGLAVEEAGHNGTLLAGAYTISKLITQKLDGLKNSERLKEKIENAKKCSEDFTKKL 317
         ||||||||||.|||||||||||||||||||.||||||||..|||||||||||||||||||
Db    70 KIQQNGLAVEAGHNGTLLAGAYTISKLITQELDGLKNSEKLKEKIENAKKCSEDFTKKL 129

Qy   318 EGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLAKAAKEMLANS 373
         |||||||||||||||||||||||||||||||||||||||||||||||||||| | |
Db   130 EGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLAKAAKETLNNS 185
```

RESULT 14
ID   Q9R7B4
AC   Q9R7B4;

PRELIMINARY;   PRT;   175 AA.

```
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C (FRAGMENT).
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TETS;
RX   MEDLINE=97478003; PubMed=9336916;
RA   Ras N.M., Postic D., Foretz M., Baranton G.;
RT   "Borrelia burgdorferi sensu stricto, a bacterial species 'made in the
RT   U.S.A.'?";
RL   Int. J. Syst. Bacteriol. 47:1112-1117(1997).
RN   [2]
SQ   SEQUENCE   175 AA;  18573 MW;  EC059E7BD3AC3250 CRC64;
FT   NON_TER       1      1
FT   NON_TER     175    175

RL   Marti-Ras N., Postic D., Foretz M., Baranton G.;
RL   Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U91792; AAB81889.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.

Query Match             44.8%;  Score 829; DB 2;  Length 175;
Best Local Similarity   97.6%;  Pred. No. 4.6e-32; Indels  0; Gaps 0;
Matches  166;  Conservative   3;  Mismatches   1;

Qy    1  MACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKA  60
         ::|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6  ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA  65

Qy   61  IGKKIHQNNGLDTENNHNGSLLAGAVAISTLIKQKLDGLKNEGLKREKIDAAKKCSETFTN  120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   66  IGKKIHQNNGLDTENNHNGSLLAGAVAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN  125

Qy  121  KLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKA  170
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  126  KLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKA  175
```

```
ID   Q44995     PRELIMINARY;   PRT;   178 AA.
AC   Q44995;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C (FRAGMENT).
GN   OSPC.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95286481; PubMed=7768799;
RA   Theisen M., Borre M., Mathiesen M.J., Mikkelsen B., Lebech A.M.,
RA   Hansen K.;
RT   "Evolution of the Borrelia burgdorferi outer surface protein OspC.";
RL   J. Bacteriol. 177:3036-3044(1995).
DR   EMBL; X84779; CAA59250.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
FT   NON_TER       1      1
FT   NON_TER     178    178
SQ   SEQUENCE   178 AA;  18911 MW;  EE50CE48EADE1CA CRC64;
```

```
Query Match             44.6%;  Score 826; DB 2;  Length 178;
Best Local Similarity   98.8%;  Pred. No. 6.5e-32; Indels  0; Gaps 0;
Matches  171;  Conservative   0;  Mismatches   2;

Qy  205  NTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDELAITKAIGKKIQONGG  264
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  NTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDELATKAIGKKIQONGG  60

Qy  265  LAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKEKIENAKKCSEDFTKKLEGEHAQL  324
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  LAVEAGHNGTSLAGAYTISKLITQKLDGLKNSEKLKEKIENAKKCSEDFTKKLEGEHAQL  120

Qy  325  GIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLAKAAKEMLANSVKEL  377
         |||||||||||||||||||||||||||||||||    ||||||||||||||||
Db  121  GIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLPKAAKEMLANSVKEL  173
```

Search completed: March 18, 2002, 10:10:54
Job time: 977 sec

THIS PAGE BLANK (USPTO)

OM protein - protein search, using sw model

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

Run on: March 18, 2002, 09:54:31 ; Search time 118.14 Seconds
                                          (without alignments)
                                          232.615 Million cell updates/sec

Title: US-09-596-746A-34
Sequence: 1 MACNNSGKDGNTSANSADES.........KNLSKAAKEMLTNSVKELTS 371
Perfect score: 1809
Searched: 522463 seqs, 74073290 residues
Scoring table: BLOSUM62
               Gapop 10.0 , Gapext 0.5

Post-processing: Minimum Match 0%
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
               Maximum Match 100%
               Listing first 45 summaries

Total number of hits satisfying chosen parameters: 522463

Database :   A_Geneseq_1101:*
               1: /SIDS8/gcgdata/geneseq/geneseqp/AA1980.DAT:*
               2: /SIDS8/gcgdata/geneseq/geneseqp/AA1981.DAT:*
               3: /SIDS8/gcgdata/geneseq/geneseqp/AA1982.DAT:*
               4: /SIDS8/gcgdata/geneseq/geneseqp/AA1983.DAT:*
               5: /SIDS8/gcgdata/geneseq/geneseqp/AA1984.DAT:*
               6: /SIDS8/gcgdata/geneseq/geneseqp/AA1985.DAT:*
               7: /SIDS8/gcgdata/geneseq/geneseqp/AA1986.DAT:*
               8: /SIDS8/gcgdata/geneseq/geneseqp/AA1987.DAT:*
               9: /SIDS8/gcgdata/geneseq/geneseqp/AA1988.DAT:*
              10: /SIDS8/gcgdata/geneseq/geneseqp/AA1989.DAT:*
              11: /SIDS8/gcgdata/geneseq/geneseqp/AA1990.DAT:*
              12: /SIDS8/gcgdata/geneseq/geneseqp/AA1991.DAT:*
              13: /SIDS8/gcgdata/geneseq/geneseqp/AA1992.DAT:*
              14: /SIDS8/gcgdata/geneseq/geneseqp/AA1993.DAT:*
              15: /SIDS8/gcgdata/geneseq/geneseqp/AA1994.DAT:*
              16: /SIDS8/gcgdata/geneseq/geneseqp/AA1995.DAT:*
              17: /SIDS8/gcgdata/geneseq/geneseqp/AA1996.DAT:*
              18: /SIDS8/gcgdata/geneseq/geneseqp/AA1997.DAT:*
              19: /SIDS8/gcgdata/geneseq/geneseqp/AA1998.DAT:*
              20: /SIDS8/gcgdata/geneseq/geneseqp/AA1999.DAT:*
              21: /SIDS8/gcgdata/geneseq/geneseqp/AA2000.DAT:*
              22: /SIDS8/gcgdata/geneseq/geneseqp/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	Query Match	% Length	DB ID	Description
1	1804	99.7	370	22	AAB62715	Borrelia sp chimer
2	1801	99.6	394	22	AAB62732	Borrelia sp chimer
3	1559.5	86.2	394	22	AAB62730	Borrelia sp chimer
4	1556.5	86.0	369	22	AAB62716	Borrelia sp chimer
5	1533.5	84.8	393	22	AAB62731	Borrelia sp chimer
5	1529	84.8	378	22	AAB62712	Borrelia sp chimer
7	1526	84.4	374	22	AAB62710	Borrelia sp chimer
8	1524.5	84.3	398	22	AAB62728	Borrelia sp chimer
9	1489	82.3	401	22	AAB62738	Borrelia sp chimer
10	1488.5	82.3	368	22	AAB62718	Borrelia sp chimer
11	1487.5	82.2	378	22	AAB62725	Borrelia sp chimer
			369	22	AAB62714	Borrelia sp chimer

12	1486	82.1	392	22	AAB62734	Borrelia sp chimer
13	1484.5	82.1	393	22	AAB62732	Borrelia sp chimer
14	1480.5	81.8	401	22	AAB62733	Borrelia sp chimer
15	1289	81.8	377	22	AAB62711	Borrelia sp chimer
16	1284.5	71.3	377	22	AAB62711	Borrelia sp chimer
17	1281.5	71.0	373	22	AAB62729	Borrelia sp chimer
18	1280	70.8	397	22	AAB62729	Borrelia sp chimer
19	1244.5	70.8	400	22	AAB62739	Borrelia sp chimer
20	1241.5	68.8	367	22	AAB62735	Borrelia sp chimer
21	1225	68.6	391	22	AAB62726	Borrelia sp chimer
22	1222	67.7	384	22	AAB62737	Borrelia sp chimer
23	1199	67.6	408	22	AAB62704	Borrelia sp chimer
24	1177.5	66.3	560	22	AAB62724	Borrelia sp chimer
25	1169.5	65.1	368	22	AAB62717	Borrelia sp chimer
26	1097	64.6	391	22	AAB62736	Borrelia sp chimer
27	1096	60.6	410	22	AAB62740	Borrelia sp chimer
28	920	60.6	386	22	AAB62727	Borrelia sp chimer
29	903	50.9	386	22	AAB62701	B. burgdorferi OspC
30	886	49.9	193	22	AAB62704	B. burgdorferi OspC
31	885	49.0	193	22	AAW41822	B. burgdorferi sen
32	884	48.9	211	18	AAB62707	B. burgdorferi OspC
33	879	48.9	192	22	AAR60888	Borrelia 297 antig
34	852	48.6	193	15	AAR60889	Borrelia 297 antig
35	837	46.3	176	15	AAR62775	Borrelia 257 antig
36	822	46.3	211	20	AAW93473	B. burgdorferi Osp
37	702	45.4	177	15	AAR62774	Borrelia 257 antig
38	701	38.8	466	16	AAR75739	B31 outer surface
39	692	38.8	587	16	AAR75746	B31 Osp-A/antigen
40	664	38.3	466	16	AAR75740	B31 outer surface
41	661.5	36.7	193	22	AAR62709	B. burgdorferi OspC
42	659.5	36.5	192	22	AAB62721	B. burgdorferi OspC
43	658.5	36.5	209	22	AAB62720	B. burgdorferi OspC
44	658.5	36.4	192	22	AAB62703	B. burgdorferi OspC
45	656.5	36.4	210	16	AAW11935	Outer surface prot
	656.5	36.3	192	15	AAR60886	Borrelia IP2 OspC

ALIGNMENTS

RESULT 1
AAB62715

ID	AAB62715 standard; protein; 370 AA.
XX	
AC	AAB62715;
XX	
DT	03-APR-2001 (first entry)
XX	
DE	Borrelia sp chimeric ospC protein SEQ ID NO: 34.
XX	
KW	Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX	
OS	Chimeric - Borrelia sp.
OS	Chimeric - Borrelia sp.
XX	
PN	WO200078966-A1.
XX	
PD	28-DEC-2000.
XX	
PF	19-JUN-2000; 2000WO-US16915.
XX	
PR	18-JUN-1999; 99US-0140042.
XX	
PA	(UNY) UNIV NEW YORK STATE RES FOUND.
PA	(BROO-) BROOK BIOTECHNOLOGIES INC.
XX	
PI	Dattwyler RJ, Seinost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
DR	WPI; 2001-050113/06.
DR	N-PSDB; AAF29019.
XX	
PT	Compositions of OspC polypeptides from strains of Borrelia which cause
PT	Lyme disease are used to immunize animals and detect immune responses
PT	to Lyme disease -

RESULT 2

AAB62730;

AC　AAB62730;

DT　03-APR-2001　(first entry)

DE　Borrelia sp chimeric ospc protein SEQ ID NO: 64.

KW　Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.

OS　Chimeric - Borrelia sp.
OS　Chimeric - Borrelia sp.

PN　WO200078966-A1.

PD　28-DEC-2000.

PF　19-JUN-2000; 2000WO-US16915.

PR　18-JUN-1999;　99US-0140042.

PA　(UYNY) UNIV NEW YORK STATE RES FOUND.
PA　(BROO-) BROOK BIOTECHNOLOGIES INC.

PI　Dattwyler RJ, Seinost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;

XX

PT　The present invention provides compositions comprising ospc proteins and
CC　chimeric ospC proteins from members of the Borrelia genus. These may be
CC　vaccines against Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC　vaccines against Borrelia infection, which is spread by ticks and leads
CC　to Lyme disease.

XX

PS　Claim 43; Page 91-92; 160pp; English.

XX

SQ　Sequence　370 AA;

Query Match　99.7%;　Score 1804;　DB 22;　Length 370;
Best Local Similarity　100.0%;　Pred. No. 1.6e-104;
Matches 370;　Conservative　0;　Mismatches　0;　Indels　0;　Gaps　0;

Qy　2　ACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDELAKAIG 61
　　　|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　1　acnnsgkdgntsansadesvkgpnlteiskkitdsnavllavkeveallssidelakaig 60

Qy　62　KKIKNDGSLDNEANRNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEEFSTK 121
　　　|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　61　kkikndgsldneanrnesllagaytistlitqklsklngseglkekiaaakkcseefstk 120

Qy　122　LKDNHAQLGIGQGVTDENAKKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANSVK 181
　　　|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　121　lkdnhaqlgigqgvtdenakkailkanaagkdkgveeleklsgsleslskaakemlansvk 180

Qy　182　ELTSPVVHGNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLTSID 241
　　　|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　181　eltspvvhgnnsgkdgntsansadesvkgpnlteiskkitesnavvlavkevetlltsid 240

Qy　242　ELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAEIEKAKKC 301
　　　|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　241　elakaigkkikndvsldneadhngslisgaylisnlitkkisaikdsgelkaeiekakkc 300

Qy　302　SEEFTAKILKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLSKAAKEM 361
　　　|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　301　seeftakilkgehtdlgkegvtddnakkailktnndktkgadeleklfesvknlskaakem 360

Qy　362　LTNSVKELTS 371
　　　||||||||||
Db　361　ltnsvkelts 370

RESULT 3

AAB62716

ID　AAB62716 standard; Protein; 369 AA.

XX

AC　AAB62716;

XX

DT　03-APR-2001　(first entry)

DE　Borrelia sp chimeric ospc protein SEQ ID NO: 36.

KW　Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.

OS　Chimeric - Borrelia sp.
OS　Chimeric - Borrelia sp.

XX

PN　WO200078966-A1.

PD　28-DEC-2000.

PF　19-JUN-2000; 2000WO-US16915.

PR　18-JUN-1999;　99US-0140042.

DR　WPI; 2001-050113/06.
DR　N-PSDB; AAF29034.

XX

PT　Compositions of. Ospc polypeptides from strains of Borrelia which cause
PT　Lyme disease are used to immunize animals and detect immune responses
PT　to Lyme disease -

XX

CC　The present invention provides compositions comprising ospc proteins and
CC　chimeric ospC proteins from members of the Borrelia genus. These may be
CC　Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC　vaccines against Borrelia infection, which is spread by ticks and leads
CC　to Lyme disease.

XX

PS　Claim 43; Page 128-129; 160pp; English.

XX

SQ　Sequence　394 AA.

Query Match　99.6%;　Score 1801;　DB 22;　Length 394;
Best Local Similarity　99.7%;　Pred. No. 2.6e-104;
Matches 369;　Conservative　1;　Mismatches　0;　Indels　0;　Gaps　0;

Qy　2　ACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDELAKAIG 61
　　　:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　25　scnnsgkdgntsansadesvkgpnlteiskkitdsnavllavkeveallssidelakaig 84

Qy　62　KKIKNDGSLDNEANRNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEEFSTK 121
　　　|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　85　kkikndgsldneanrnesllagaytistlitqklsklngseglkekiaaakkcseefstk 144

Qy　122　LKDNHAQLGIGQGVTDENAKKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANSVK 181
　　　|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　145　lkdnhaqlgigqgvtdenakkailkanaagkdkgveeleklsgsleslskaakemlansvk 204

Qy　182　ELTSPVVHGNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLTSID 241
　　　|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　205　eltspvvhgnnsgkdgntsansadesvkgpnlteiskkitesnavvlavkevetlltsid 264

Qy　242　ELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAEIEKAKKC 301
　　　|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　265　elakaigkkikndvsldneadhngslisgaylisnlitkkisaikdsgelkaeiekakkc 324

Qy　302　SEEFTAKILKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLSKAAKEM 361
　　　|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　325　seeftakilkgehtdlgkegvtddnakkailktnndktkgadeleklfesvknlskaakem 384

Qy　362　LTNSVKELTS 371
　　　||||||||||
Db　385　ltnsvkelts 394

RESULT 4

ID   AAB62731 standard; Protein; 393 AA.
XX
AC   AAB62731;
XX
DT   03-APR-2001 (first entry)
XX
DE   Borrelia sp chimeric ospC protein SEQ ID NO: 66.
XX
KW   Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
XX
PN   WO200078966-A1.

XX
PD   28-DEC-2000.
XX
PF   19-JUN-2000; 2000WO-US016915.
XX
PR   18-JUN-1999; 99US-0140042.
XX
PA   (UYNY) UNIV NEW YORK STATE RES FOUND.
PA   (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI   Dattwyler RJ, Seinost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
DR   WPI; 2001-050113/06.
DR   N-PSDB; AAF29035.
XX
PT   The present invention provides compositions comprising ospC proteins and
PT   chimeric ospC proteins from members of the Borrelia genus. These may be
CC   Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC   vaccines against Borrelia infection, which is spread by ticks and leads
CC   to Lyme disease.
XX
PS   Claim 43; Page 131; 160pp; English.
XX
SQ   Sequence   393 AA;

Query Match   86.0%; Score 1556.5; DB 22; Length 393;
Best Local Similarity   87.3%; Pred. No. 3.7e-89;
Matches 322; Conservative 21; Mismatches 25; Indels   1; Gaps   1;

Qy   2 ACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDELAKAIG 61
          :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 ACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDELAKAIG 61

Qy   62 KKIKNDGSLDNEANRNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEEFSTK 121
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   25 scnnsgkdgntsansadesvkgpnlteiskkitdsnavllavkeveallssidelakaig 84

Qy   122 LKDNHAOLGIQGVTDENAKKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANSVK 181
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   85 kkikndgsldneanrnesllagaytistlitqklsklngseglkekiaaakkcseefstk 144

Qy   182 ELTSPVVHGNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLTSID 241
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 lkdnhaqlgiqgvtdenakkailkanaagkdkgveelekl sgslesl skaakemlansvk 181

Qy   242 ELA-KAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAEIEKAKK 300
          ||| ||||||||:|||| ||| ||| ||:| ||::: :|:| |||||||||| ||
Db   182 ELTSPVVHGNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLTSID 241

Qy   301 CSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLSKAAKE 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   205 eltspvvhgnnsgkdgntsansadesvkgpnlteiskkitesnavvlavkevetlltsid 264

Qy   361 MLTNSVKEL 369
          ||||||||
Db   265 elatkaigkkiqqngglaveaghngtllagaytiskiltqkidglknseklkekienakk 324

Qy   301 CSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLSKAAKE 360
          |||:|:| || |||| :| | ::| |||| ||::| ||:|||||::||:||||| |
Db   325 csedftkklegehaqlgienvtdenakkaillltdaakdkgaaeleklfkavenlakaake 384

Qy   361 MLTNSVKEL 369
          || |||||
Db   385 mlansvkel 393

Qy   361 MLTNSVKEL 369
          ||||||||
Db   361 mlansvkel 369

RESULT 5

ID   AAB62712 standard; Protein; 378 AA.
XX
AC   AAB62712;
XX
DT   03-APR-2001 (first entry)
XX
DE   Borrelia sp chimeric ospC protein SEQ ID NO: 28.
XX

```
KW   Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
XX
PN   WO200078966-A1.
XX
PD   28-DEC-2000.
XX
PF   19-JUN-2000; 2000WO-US16915.
XX
PR   18-JUN-1999; 99US-0140042.
XX
PA   (UYNY ) UNIV NEW YORK STATE RES FOUND.
PA   (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI   Dattwyler RJ, Seinost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
PR   WPI; 2001-050113/06.
     N-PSDB; AAF29016.
XX
CC   The present invention provides compositions comprising ospc proteins and
CC   chimeric ospc proteins from members of the Borrelia genus. These may be
CC   Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC   vaccines against Borrelia infection, which is spread by ticks and leads
CC   to Lyme disease.
XX
PT   Compositions of OspC polypeptides from strains of Borrelia which cause
PT   Lyme disease are used to immunize animals and detect immune responses
PT   to Lyme disease -
XX
PS   Claim 43; Page 83-84; 160pp; English.
XX
SQ   Sequence    378 AA;

Query Match    84.8%;  Score 1533.5;  DB 22;  Length 378;
Best Local Similarity   86.3%;  Pred. No. 9.5e-88;
Matches  328;  Conservative 10;  Mismatches  31;  Indels  11;  Gaps  4;

Qy    1 MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEL-AKA 59
        |||||||||||||||||||||||||||  ||||||||||||||||||||||||||| |||
Db    1 macnnsgkdgntsansadesvkgpnlteinkkitdsnavllavkeveallssideiaaka 60

Qy   60 IGKKINDGSLDNEANRNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEEFS 119
        ||||    :         |||||||| :||||:||  |||| :||||||:|||||| !:
Db   61 igkkihgnngldtennhngsllagayaistlikqkldglk-neglkekidaakkcsetft 119

Qy  120 TKLKDNHAQLGIQGVTDENAKKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANS 179
        ||||||||||||||||||:|||||||||::  ||:| :|||:|| ||:|||||||||||
Db  120 nklkekhtdlgkegvtdadakeailkangt-ktkgaeelgklfesvevlskaakemlans 178

Qy  180 VKELTSPVVHG------NNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVK 231
        |||||||||       |||||||||||||||||||||||||||||||||||||||||
Db  179 vkeltspvvaespkkpsmvnnsgkdgntsansadesvkgpnlteiskkitesnavvlavk 238

Qy  232 EVETLLTSIDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNIITKKISAIKDSGEL 291
        ||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||
Db  239 evetlltsidelakaigkkikndvslneadhngslisgaylisnlitkkisaikdsgel 298

Qy  292 KAEIEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESV 351
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  299 kaeiekakkcseeftaklkgehtdlgkegvtddnakkailktnndktkgadeleklfesv 358

Qy  352 KNLSKAAKEMLTNSVKELTS 371
        ||||||||||||||||||||
Db  359 knlskaakemltnsvkelts 378
```

```
AC   AAB62710;
XX
DT   03-APR-2001  (first entry)
XX
DE   Borrelia sp chimeric ospc protein SEQ ID NO: 24.
XX
KW   Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
XX
PN   WO200078966-A1.
XX
PD   28-DEC-2000.
XX
PF   19-JUN-2000; 2000WO-US16915.
XX
PR   18-JUN-1999; 99US-0140042.
XX
PA   (UYNY ) UNIV NEW YORK STATE RES FOUND.
PA   (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI   Dattwyler RJ, Seinost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
DR   WPI; 2001-050113/06.
DR   N-PSDB; AAF29014.
XX
CC   The present invention provides compositions comprising ospc proteins and
CC   chimeric ospc proteins from members of the Borrelia genus. These may be
CC   Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC   vaccines against Borrelia infection, which is spread by ticks and leads
CC   to Lyme disease.
XX
PT   Compositions of OspC polypeptides from strains of Borrelia which cause
PT   Lyme disease are used to immunize animals and detect immune responses
PT   to Lyme disease -
XX
PS   Claim 43; Page 78-79; 160pp; English.
XX
SQ   Sequence    374 AA;

Query Match    84.5%;  Score 1529;  DB 22;  Length 374;
Best Local Similarity   87.0%;  Pred. No. 1.8e-87;
Matches  327;  Conservative  9;  Mismatches  32;  Indels  8;  Gaps  4;

Qy    2 ACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEL-AKAI 60
        ||||||||||||||||||||||||||  ||||||||||||||||||||||||||| ||||
Db    1 acnnsgkdgntsansadesvkgpnlteinkkitdsnavllavkeveallssideiaakai 60

Qy   61 GKKINDGSLDNEANRNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEEFST 120
        ||||    :         |||||||| :||||:||  |||| :||||||:|||||| !:
Db   61 gkkihgnngldtennhngsllagayaistlikqkldglk-neglkekidaakkcsetftn 119

Qy  121 KLKDNHAQLGIQGVTDENAKKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANSV 180
        |||||||||||||||||:|||||||||::  ||:| :|||:|| ||:||||||||||||
Db  120 klkekhtdlgkegvtdadakeailkangt-ktkgaeelgklfesvevlskaakemlansv 178

Qy  181 KELTSPVVHG------NNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVET 235
        ||||||||       ||||||||||||||||||||||||||||||||||||||||||||
Db  179 keltspvvaespanvnnsgkdgntsansadesvkgpnlteiskkitesnavvlavkevet 238

Qy  236 LLTSIDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNIITKKISAIKDSGELKAEI 295
        |||||||||||||||||||||||| ||||||||||||||||:|||||||||||||||||
Db  239 lltsidelakaigkkikndvslneadhngslisgaylisnlitkkisaikdsgelkaei 298

Qy  296 EKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLS 355
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  299 ekakkcseeftaklkgehtdlgkegvtddnakkailktnndktkgadeleklfesvknls 358

Qy  356 KAAKEMLTNSVKELTS 371
        ||||||||||||||||
Db  359 kaakemltnsvkelts 374
```

```
RESULT  6
AAB62710
ID   AAB62710 standard; Protein; 374 AA.
XX
```

```
RESULT  7

AAB62728
ID    AAB62728 standard; Protein; 398 AA.
XX
AC    AAB62728;
XX
DT    03-APR-2001  (first entry)
XX
DE    Borrelia sp chimeric ospC protein SEQ ID NO: 60.
XX
KW    Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS    Chimeric - Borrelia sp.
OS    Chimeric - Borrelia sp.
XX
PN    WO200078966-A1.
XX
PD    28-DEC-2000.
XX
PN    28-DEC-2000.
XX
OS    19-JUN-2000; 2000WO-US16915.
XX
PR    18-JUN-1999;    99US-0140042.
XX
PA    (UYNY ) UNIV NEW YORK STATE RES FOUND.
PA    (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI    Dattwyler RJ,  Seinost G,  Dykhuizen D,  Luft BJ,  Gomes-Solecki M;
XX
DR    WPI; 2001-050113/06.
DR    N-PSDB; AAF29032.
XX
PT    Compositions of OspC polypeptides from strains of Borrelia which cause
PT    Lyme disease are used to immunize animals and detect immune responses
PT    to Lyme disease -
XX
CC    Claim 43: Page 123; 160pp; English.
XX
CC    The present invention provides compositions comprising ospC proteins and
CC    chimeric ospC proteins from members of the Borrelia genus. These may be
CC    Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC    vaccines against Borrelia infection, which is spread by ticks and leads
CC    to Lyme disease.
XX
SQ    Sequence     398 AA;

Query Match             84.4%;  Score 1526;  DB 22;  Length 398;
Best Local Similarity   86.7%;  Pred. No. 2.9e-87;
Matches 326;  Conservative  10;  Mismatches  32;  Indels   8;  Gaps   4;

Qy      2  ACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEL-AKAI   60
           |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db     25  scnnsgkdgntsansadesvkgpnlteiskkitdsnavllavkeveallssideiaakai  84

Qy     61  GKKIKNDGSLDNEANRNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEEFST  120
           ||  |||| |||||||||||||| ||||||||||||||||||||||||| ||||||||:
Db     85  gkkihqnngldteynhngsllagayaistlikqkldgik-neglkekidaakkcsetftn  143

Qy    121  KLKDNHAQLGIQGVTDENAKKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANSV  180
           |||:| |||||:|:||||||||:|:|||||||||||||||||||||||||||||||
Db    144  klkdnhtdlgkegvtdadakeailktngt-ktkgaeelgklfesvevlskaakemlansv  202

Qy    181  KELTSPVVHG-----NNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVET  235
           ||||||||||     |||||||||||||||||||||||||||||||||||||||||||
Db    203  keltspvvaespamvnnsgkdgntsansadesvkgpnlteiskkitesnavvlavkevet  262

Qy    236  LLTSIDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAEI  295
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    263  litsidelakaigkkikndvsldneadhngslisgaylisnlitkkisaikdsgelkaei  322

Qy    296  ERAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLS  355
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    323  ekakkcseeftaklkgehtdlgkegvtddnakkailktnndktkgadeleklfesvknls  382

Qy    356  KAAAKEMLTNSVKELTS  371
           |||||||||||||||||
Db    383  kaakemltnsvkelts   398
```

```
RESULT  8

AAB62738
ID    AAB62738 standard; Protein; 401 AA.
XX
AC    AAB62738;
XX
DT    03-APR-2001  (first entry)
XX
DE    Borrelia sp chimeric ospC protein SEQ ID NO: 80.
XX
KW    Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS    Chimeric - Borrelia sp.
OS    Chimeric - Borrelia sp.
XX
PN    WO200078966-A1.
XX
PD    28-DEC-2000.
XX
OS    19-JUN-2000; 2000WO-US16915.
XX
PR    18-JUN-1999;    99US-0140042.
XX
PA    (UYNY ) UNIV NEW YORK STATE RES FOUND.
PA    (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI    Dattwyler RJ,  Seinost G,  Dykhuizen D,  Luft BJ,  Gomes-Solecki M;
XX
DR    WPI; 2001-050113/06.
DR    N-PSDB; AAF29042.
XX
PT    Compositions of OspC polypeptides from strains of Borrelia which cause
PT    Lyme disease are used to immunize animals and detect immune responses
PT    to Lyme disease -
XX
CC    Claim 43: Page 150-151; 160pp; English.
XX
CC    The present invention provides compositions comprising ospC proteins and
CC    chimeric ospC proteins from members of the Borrelia genus. These may be
CC    Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC    vaccines against Borrelia infection, which is spread by ticks and leads
CC    to Lyme disease.
XX
SQ    Sequence     401 AA;

Query Match             84.3%;  Score 1524.5;  DB 22;  Length 401;
Best Local Similarity   86.0%;  Pred. No. 3.7e-87;
Matches 326;  Conservative  10;  Mismatches  32;  Indels  11;  Gaps   4;

Qy      2  ACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEL-AKAI   60
           |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db     25  scnnsgkdgntsansadesvkgpnlteiskkitdsnavllavkeveallssideiaakai  84

Qy     61  GKKIKNDGSLDNEANRNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEEFST  120
           ||  |||| |||||||||||||| ||||||||||||||||||||||||| ||||||||:
Db     85  gkkihqnngldteynhngsllagayaistlikqkldgik-neglkekidaakkcsetftn  143

Qy    121  KLKDNHAQLGIQGVTDENAKKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANSV  180
           |||:| |||||:|:||||||||:|:|||||||||||||||||||||||||||||||
Db    144  klkdnhtdlgkegvtdadakeailktngt-ktkgaeelgklfesvevlskaakemlansv  202

Qy    181  KELTSPVVHG------NNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKE   232
           ||||||||||      |||||||||||||||||||||||||||||||||||||||||
Db    203  keltspvvaespkpsmvnnsgkdgntsansadesvkgpnlteiskkitesnavvlavke  262
```

Left column (RESULT 9)

```
Qy  233 VETLLTSIDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELK 292
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  263 vetlltsidelakaigkkikndvsldneadhngslisgaylisnlitkkisaikdsgelk 292

Qy  293 AEIEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKAILKTNNDKTKGADELEKLFESVK 352
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  323 aeiekakkcseeftaklkgehtdlgkegvtddnakailktnndktkgadeleklfesvk 382

Qy  353 NLSKAAKEMLTNSVKELTS 371
        |||||||||||||||||||
Db  383 nlskaakemltnsvkelts 401
```

```
RESULT   9
AAB62718

ID   AAB62718 standard; Protein; 368 AA.
XX
DT   03-APR-2001   (first entry)
XX
DE   Borrelia sp chimeric ospc protein SEQ ID NO: 40.
XX
KW   Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
XX
PN   WO200078966-A1.
XX
PD   28-DEC-2000.
XX
PF   19-JUN-2000; 2000WO-US16915.
XX
PR   18-JUN-1999;   99US-0140042.
XX
PA   (UYNY ) UNIV NEW YORK STATE RES FOUND.
PA   (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI   Dattwyler RJ, Seinost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
DR   WPI; 2001-050113/06.
DR   N-PSDB; AAF29022.
XX
PT   Compositions of OspC polypeptides from strains of Borrelia which cause
PT   Lyme disease are used to immunize animals and detect immune responses
PT   to Lyme disease.
CC   The present invention provides compositions comprising ospC proteins and
CC   chimeric ospC proteins from members of the Borrelia genus. These may be
CC   Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC   vaccines against Borrelia infection, which is spread by ticks and leads
CC   to Lyme disease.
XX
PS   Claim 43; Page 99-100; 160pp; English.
XX
SQ   Sequence   368 AA;
```

```
Query Match             82.3%;  Score 1489;  DB 22;  Length 368;
Best Local Similarity   83.8%;  Pred. No. 5.3e-85;
Matches  311;  Conservative   25;  Mismatches   31;  Indels    4;  Gaps    4;
```

```
Qy    2 ACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDELA-KAI 60
        |||||||| | ||||||||||||||||||||||| |||| :|||||||| ||||||| ||
Db    1 acnnsgkdgnasansadesvkgpnlteiskkitesnavvlavkevetllasidelatkai 60

Qy   61 GKKIKNDGSLDNEANRNESLLAGAYTISTLITQKLSKLINGSEGLKEKIAAAKKCSEEFST 120
        |||| |: : :||  : ||| ||| :|| ||::|| |:|| ||||||| : ||||||||:
Db   61 gkkigng-leangskntslisgayaisdliaekinvlknee-lkekidtakqcsteftn 118

Qy  121 KLKDNHAQLGIQGYTDENAKKAILKANAAGKDKGVEELEKLSGGSLESLSKAAKEMLANSV 180
```

Right column (RESULT 10)

```
Db  119 klksehavlgldnltddnaqrailkkh-ankdkgaaaeleklfkavenlskaaqdtlknav 177

Qy  181 KELTSPVVHGNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLTSI 240
        ||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  178 keltspivhgnnsgkdgntsansadesvkgpnlteiskkitesnavvlavkevetlltsi 237

Qy  241 DELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAEIEKAKK 300
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  238 delakaigkkikndvsldneadhngslisgaylisnlitkkisaikdsgelkaeiekakk 297

Qy  301 CSEEFTAKLKGEHTDLGKEGVTDDNAKAILKTNNDKTKGADELEKLFESVKNLSKAAKE 360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  298 cseeftaklkgehtdlgkegvtddnakailktnndktkgadeleklfesvknlskaake 357

Qy  361 MLTNSVKELTS 371
        |||||||||||
Db  358 mltnsvkelts 368
```

```
RESULT  10
AAB62725

ID   AAB62725 standard; Protein; 378 AA.
XX
AC   AAB62725;
XX
DT   03-APR-2001   (first entry)
XX
DE   Borrelia sp chimeric ospc protein SEQ ID NO: 54.
XX
KW   Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
XX
PN   WO200078966-A1.
XX
PD   28-DEC-2000.
XX
PF   19-JUN-2000; 2000WO-US16915.
XX
PR   18-JUN-1999;   99US-0140042.
XX
PA   (UYNY ) UNIV NEW YORK STATE RES FOUND.
PA   (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI   Dattwyler RJ, Seinost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
DR   WPI; 2001-050113/06.
DR   N-PSDB; AAF29029.
XX
PT   Compositions of OspC polypeptides from strains of Borrelia which cause
PT   Lyme disease are used to immunize animals and detect immune responses
PT   to Lyme disease.
CC   The present invention provides compositions comprising ospC proteins and
CC   chimeric ospC proteins from members of the Borrelia genus. These may be
CC   Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC   vaccines against Borrelia infection, which is spread by ticks and leads
CC   to Lyme disease.
XX
PS   Claim 43; Page 114-115; 160pp; English.
XX
SQ   Sequence   378 AA;
```

```
Query Match             82.3%;  Score 1488.5;  DB 22;  Length 378;
Best Local Similarity   83.3%;  Pred. No. 5.8e-85;
Matches  309;  Conservative   31;  Mismatches   28;  Indels    3;  Gaps    2;
```

```
Qy    1 MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDELAKAI 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 macnnsgkdgntsansadesvkgpnlteiskkitdsnavllavkeveallssidelakai 60
```

RESULT 11

```
ID   AAB62714 standard; Protein; 369 AA.
XX
AC   AAB62714;
XX
DT   03-APR-2001   (first entry)
XX
DE   Borrelia sp chimeric ospc protein SEQ ID NO: 32.
XX
KW   Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
OS   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
XX
PN   WO200078966-A1.
XX
PD   28-DEC-2000.
PF   19-JUN-2000; 2000WO-US16915.
PR   18-JUN-1999;   99US-0140042.
XX
PI   Dattwyler RJ, Selnost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
     (UYNY ) UNIV NEW YORK STATE RES FOUND.
     (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
DR   WPI; 2001-050113/06.
DR   N-PSDB; AAF29018.
XX
PS   Claim 43; Page 89; 160pp; English.
XX
PT   Compositions of OspC polypeptides from strains of Borrelia which cause
PT   Lyme disease are used to immunize animals and detect immune responses
PT   to Lyme disease -
XX
CC   The present invention provides compositions comprising ospc proteins and
CC   chimeric ospc proteins from members of the Borrelia genus. These may be
CC   Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC   vaccines against Borrelia infection, which is spread by ticks and leads
CC   to Lyme disease.
XX
SQ   Sequence 369 AA;

Query Match             82.2%;   Score 1487.5;   DB 22;   Length 369;

Qy     61 GKRIKNDGSLDNEANRNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEEFST 120
           ||||||||||||||||||||||||| ||||:||||||| :||||||||||||||||||||
Db     61 gkkikndgsldneanrnesllagaytistliitqklskingseglkekiaaakkcseefst 120

Qy    121 KLKDNHAQLGIQGVTDENAKKAILKANAAGKDKGVEELEKLSGGSLESLSKAAKEMLANSV 180
           ||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db    121 klkdnhaqlgiqgvtdenakkailkanaagkdkgveeleklsgsliesl skaakemlansv 180

Qy    181 KELTSPVVHGNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLTSI 240
           |||||||||||||:   :::|| : ::||  |||| |||||:|||:|||||||||| :||
Db    181 keltspvvhgnnsg--gdsastnpdesakgpnlteiskkitdsnafllavkeveallssi 238

Qy    241 DELAKAIGKKIKNDVSLDNEADHNGSLISGAVLISNLITKKISAIKDSGELKAEIEKAKK 300
           ||| :|||||||||  :||| |:|| ||:||| ::|||:|| : |:|  |:|  | ||||
Db    239 delaskaigkkikndgtidneanrnesllagayeiskiitqklsvl-nseelkkikeakd 297

Qy    301 CSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLSKAAKE 360
           || :||:| | | ||||::||:|||||||||||:|  :|||| |:| ||:|:||||||::
Db    298 csqkftkikdshaelgiqsvgqddnakkailkthgtkdkgakeleelfksleslskaaqa 357

Qy    361 MLTNSVKELTS 371
           ||||||||||
Db    358 altnsvkeltn 368
```

RESULT 12

```
Best Local Similarity   84.4%;  Pred. No. 6.5e-85;
Matches  313;  Conservative  24;  Mismatches  31;  Indels  3;  Gaps  3;

ID   AAB62734 standard; Protein; 392 AA.
XX
AC   AAB62734;
XX
DT   03-APR-2001   (first entry)
XX
DE   Borrelia sp chimeric ospc protein SEQ ID NO: 72.
XX
KW   Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
OS   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
XX
PN   WO200078966-A1.
XX
PD   28-DEC-2000.
PF   19-JUN-2000; 2000WO-US16915.
PR   18-JUN-1999;   99US-0140042.
XX
     (UYNY ) UNIV NEW YORK STATE RES FOUND.
     (BROO-) BROOK BIOTECHNOLOGIES INC.
PI   Dattwyler RJ, Selnost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
DR   N-PSDB; AAF29038.
DR   WPI; 2001-050113/06.
XX
PS   Claim 43; Page 139-140; 160pp; English.
XX
PT   Compositions of OspC polypeptides from strains of Borrelia which cause
PT   Lyme disease are used to immunize animals and detect immune responses
PT   to Lyme disease -
XX
CC   The present invention provides compositions comprising ospc proteins and
CC   chimeric ospc proteins from members of the Borrelia genus. These may be
CC   Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC   vaccines against Borrelia infection, which is spread by ticks and leads

Qy      2 ACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDELAKAIG 61
           | ||:  |:|:|||||||||:|||||| :||||:|||:|||||||||| :||||:|||||
Db      1 acnnsgkdgntsansadesvkgpnlteiskkitdsnavllavkeveallssidelakaig 61

Qy     62 KKIKNDGSLDNEANRNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEEFSTK 121
           ||||||||||||||||||||||||| ||||:|||||| :|||||||||||||||||||||
Db     61 kkikndgsldneanrnesllagaytistliitqklskingseglkekiaaakkcseefstk 120

Qy    122 LKDNHAQLGIQGVTDENAKKAILKANAAGKDKGVEELEKLSGGSLESLSKAAKEMLANSV 181
           ||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db    121 lkdnhaqlgiqgvtdenakkailkanaagkdkgveeleklsgsliesl skaakemlansv 180

Qy    182 KELTSPVVHGNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLTSID 241
           : |||||||||| :|||| :||||||||||||||||||||||||||||||||||| :|||
Db    181 eltspvvhgnnsrkdgnastnsadesvkgpnlteiskkitesnavvlavkevetllssid 240

Qy    242 ELA-KAIGKKIKNDVSLDNEADHNGSLISGAVLISNLITKKISAIKDSGELKAEIEKAKK 300
           ||: ||||||| :|:  : :|:|:| | |||| ||: : :|:  ||:  :|:| || |::
Db    241 elatkaigkkignn-gleangsknts11sgayaisd11aekinv1kne-e1kekidtakq 298

Qy    301 CSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLSKAAKE 360
           || ||| || ||| || ::|:|||||| ||||: |:|||| | |:||| |:||||||:
Db    299 cseftnklksehavlgldnltddnagrailkkhankdkgaaelekifkavenlskaaqd 358

Qy    361 MLTNSVKELTS 371
           ||:|:|||||
Db    359 tlknavkelts 369
```

CC   to Lyme disease.
XX
SQ   Sequence   392 AA;

Qy   2 ACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDELA-KAI 60
         :|||||||| ||||||||||||||||||||||:|||||:||||||||:||:|||||| ||:
Db   25 scnnsgkdgnasansadesvkgpnlteiskkitesnavvlavkevetllasidelatkai 84

Qy   61 GKKIKNDGSLDNEANRNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEEFST 120
         || |: ||| |||||||||| |||| |:| | |:| | ||||| |::||| |:||:
Db   85 gkkignng-leangskntslisgayaisdliaeklnvlknee-lkekidtakqcsteftn 142

Qy   121 KLKDNHAQLGIQGVTDENARKKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANSV 180
         ||| :||::|: |::||| :||||:|| ||||| ||||||| || ||::|||:| :|||
Db   143 klksehavlgldnitddnagrailkkh-ankdkgaaeleklfkavenlskaaqdtiknav 201

Qy   181 KELTSPVVGHGNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLTSI 240
         ||||||:|:|||||||||||||||||||||||||||||||||||||||||||||||||||
Db   202 keltspivhgnnsgkdgntsansadesvkgpnlteiskkitesnavvlavkevetlltsi 261

Qy   241 DELAKAIGKKIKNDVSLDNEADHINGSLISGAYLISNLITKKISAIKDSGELKAEIEKAKK 300
         |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db   262 delakaigkkikndvsldneadhngslisgayiisnlitkkisaikdsgelkaeiekakk 321

Qy   301 CSEEFTAKILKGEHTDLGKEGVTDDNAKKAILKTNNDKDTKGADELEKLFESVKNLSKAAKE 360
         ||||||||||||||||||||||||||||||||||||||: |||||||||||||||||||||
Db   322 cseeftakilkgehtdlgkegvtddnakkaiktnndktkgadeleklfesvknlskaake 381

Qy   361 MLTNSVKELTS 371
         |||||||||||
Db   382 mltnsvkelts 392

RESULT 13
ID   AAB62732 standard; Protein; 393 AA.
XX
AC   AAB62732;
XX
DT   03-APR-2001 (first entry)
XX
DE   Borrelia sp chimeric ospc protein SEQ ID NO: 68.
XX
KW   Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
XX
PN   WO200078966-A1.
XX
PD   28-DEC-2000.
XX
PF   19-JUN-2000; 2000WO-US16915.
XX
PR   18-JUN-1999; 99US-0140042.
XX
PA   (UYNY) UNIV NEW YORK STATE RES FOUND.
PA   (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI   Dattwyler RJ, Seinost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
DR   WPI; 2001-050113/06.
DR   N-PSDB; AAF29036.
XX
CC   Compositions of OspC polypeptides from strains of Borrelia which cause
CC   Lyme disease are used to immunize animals and detect immune responses
CC   to Lyme disease -

SQ   Sequence   393 AA;

Query Match 82.1%; Score 1484.5; DB 22; Length 393;
Best Local Similarity 84.1%; Pred. No. 1.1e-84;
Matches 312; Conservative 25; Mismatches 31; Indels 3; Gaps 3;

Qy   2 ACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDELA-KAIG 61
         :||||||||||||||||||||||||||||||||:|||||:||||||||:||:||||| ||||
Db   25 scnnsgkdgntsansadesvkgpnlteiskkitesnavvlavkevetllasidelakaig 84

Qy   62 KKIKNDGSLDNEANRNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEEFSTK 121
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   85 kkikndgsldneanrnesllagaytistlitqklsklngseglkekiaaakkcseefstk 144

Qy   122 LKDNHAQLGIQGVTDENARKKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANSVK 181
         ||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||
Db   145 lkdnhaqlgiqgvtdenakkallkanaagkdkgveeleklsgsleslskaakemlansvk 204

Qy   182 ELTSPVVHGNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLTSID 241
         ||||||||||||:||||||:|||||||||||||||||||||||||||||||||||||||||
Db   205 eltspvvhgnnsrkdgnastnsadesvkgpnlteiskkitesnavvlavkevetlltsid 264

Qy   242 ELA-KAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAEIEKAKK 300
         ||: ||||||| | | |: | |||||||||| |:|| :|:|:| : ||:||:| ||:|
Db   265 elatkaigkkignn-gleangskntslisgayaisdliaeklnvlkne-elkekidtakq 322

Qy   301 CSEEFTAKILKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLSKAAKE 360
         ||||||||||:|||:|:|:| |:|||| ||||:| ||||| ||||| |:||||||||||:
Db   323 cseeftakilksehavigldnitddnagrailkkhankdkgaaeleklfkavenlskaaqd 382

Qy   361 MLTNSVKELTS 371
         |||:|||||||
Db   383 tlknavkelts 393

RESULT 14
ID   AAB62733 standard; Protein; 401 AA.
XX
AC   AAB62733;
XX
DT   03-APR-2001 (first entry)
XX
DE   Borrelia sp chimeric ospc protein SEQ ID NO: 70.
XX
KW   Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
XX
PN   WO200078966-A1.
XX
PD   28-DEC-2000.
XX
PF   19-JUN-2000; 2000WO-US16915.
XX
PR   18-JUN-1999; 99US-0140042.
XX
PA   (UYNY) UNIV NEW YORK STATE RES FOUND.
PA   (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI   Dattwyler RJ, Seinost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
DR   WPI; 2001-050113/06.

DR   N-PSDB; AAF29037.
XX
PT   Lyme disease are used to immunize animals and detect immune responses
PT   to Lyme disease -
XX
PS
PT
XX
PS   Claim 43; Page 136-137; 160pp; English.
XX
CC   The present invention provides compositions comprising ospc proteins and
CC   chimeric ospc proteins from members of the Borrelia genus. These may be
CC   Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC   vaccines against Borrelia infection, which is spread by ticks and leads
CC   to Lyme disease.
XX
SQ   Sequence   401 AA;

Query Match   81.8%; Score 1480.5; DB 22; Length 401;
Best Local Similarity   83.0%; Pred. No. 2e-84;
Matches   307; Conservative   32; Mismatches   28; Indels   3; Gaps   2;

Qy   2 ACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDELAKAIG   61
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   25 scnnsgkdgntsansadesvkgpnlteiskkitdsnavllavkeveallssidelakaig   84

Qy   62 KKIKNDGSLDNEANRNESLLAGAYTISTLTIQKLSKLNGSEGLKEKIAAAKKCSEEFSTK 121
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   85 kkikndgsldneanrnesllagaytistltiqklsklngseglkekiaaakkcseeefstk 144

Qy 122 LKDNHAQLGIQGVTDENAKKAILKANNAGKDKGVEELEKLSGSLESLSKAAKEMLANSVK 181
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 145 lkdnhaqlgiqgvtdenakkailkannagkdkgveeleklsgsleslskaakemlansvk 204

Qy 182 ELTSPVVHGNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLTSID 241
         ||:::|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 205 eltspvvhgnneg--gdsastnpdesakgpnltviskklitqklsvl-nseelkkikeakdc 262

Qy 242 ELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLLTKKISAIKDSGELKAEIEKAKKC 301
         |||||||||||||::|:|||||:|:||||:|||||:|||||:|||:|:||
Db 263 elakaigkkikndtgldneanrnesllagayeiskllitqklsvl-nseelkkikeakdc 321

Qy 302 SEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLSKAAKEM 361
         :|||||||||:|:|||||||||||||||||||::||||:|||||||||:|:|||||||||:
Db 322 sqkfttklkdshaelqsvqddnakkailkthgtckdkgakeleelfksleslskaaqaa 381

Qy 362 LTNSVKELTS 371
         ||||||||||:
Db 382 ltnsvkeltn 391

PA   (UYNY) UNIV NEW YORK STATE RES FOUND.
PA   (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI   Dattwyler RJ, Selnost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
PN   WPI; 2001-050113/06.
DR   N-PSDB; AAF29017.
XX
PT   Lyme disease are used to immunize animals and detect immune responses
PT   to Lyme disease -
DR   Compositions of OspC polypeptides from strains of Borrelia which cause
XX
PS   Claim 43; Page 86-87; 160pp; English.
XX
CC   The present invention provides compositions comprising ospc proteins and
CC   chimeric ospc proteins from members of the Borrelia genus. These may be
CC   Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC   vaccines against Borrelia infection, which is spread by ticks and leads
CC   to Lyme disease.
XX
SQ   Sequence   377 AA;

Query Match   71.3%; Score 1289; DB 22; Length 377;
Best Local Similarity   74.1%; Pred. No. 1.4e-72;
Matches   281; Conservative   30; Mismatches   56; Indels   12; Gaps   5;

Qy   1 MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEL-AKA   59
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||
Db   1 macnnsgkdgntsansadesvkgpnlteiskkitdsnavllavkeveallssidel-aka   60

Qy   60 IGKKIKNDGSLDNEANRNESLLAGAYTISTLTIQKLSKLNGSEGLKEKIAAAKKCSEEFS 119
         ||||:|| :|||| |||||||||||:|||||:|||| || :||||||:|||||||:|||:
Db   61 igkkihqngldtennhngsllagayaistlikqkldglk-neglkekidaakkcsetft 119

Qy 120 TKLKDNHAQLGIQGVTDENAKKAILKANNAGKDKGVEELEKLSGSLESLSKAAKEMLANS 179
         .|||:|||:|||:||:|:||||| ||.|:|||:||:|||| |:||:|||||||||||||
Db 120 nklkekhtdlgkegvcdadkealikangt-ktkgaeelgkifesvevlskaakemlans 178

Qy 180 VKELTSPVVHG------NNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVK 231
         ||||||||| | |||||||||||||||||||||||||||||||||||||||
Db 179 vkeltspvvaespkpkpsmvnnsgkdgntsansadesvkgpnlteiskkitesnavvlavk 238

Qy 232 EVETLLTSIDELA-KAIGKKIKNDVSLDNEADHNGSLISGAYLISNLLTKKISAIKDSGE 290
         |:|:||:||||| ||||||:|:| ::| :|:|||:||||||:||| :|||:|:| .:
Db 239 eietllasidelatkaigkkiqnggglaveaaghngtllagaytiskllitqkldgknsek 298

Qy 291 LKAEIEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFES 350
         |||:||:||||||:||||||||||:|||||||:||||||||| :|:||||||:||||||:
Db 299 lkeklenakkcsedftkklegehaqlgienvtdenakkailitdaakdkgaaeleklfka 358

Qy 351 VKNLSKAAKEMLTNSVKEL 369
         ||||:|||||||:||||||
Db 359 venlakaakemlansvkel 377

Search completed: March 18, 2002, 09:54:32
Job time: 330 sec

RESULT 15
ID   AAB62713 standard; Protein; 377 AA.
XX
AC   AAB62713;
XX
DT   03-APR-2001 (first entry)
XX
DE   Borrelia sp chimeric ospc protein SEQ ID NO: 30.
XX
OS   Borrelia sp.
OS   Chimeric - Borrelia sp.
XX
KW   Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
PN   WO200078966-A1.
XX
PD   28-DEC-2000.
XX
PF   19-JUN-2000; 2000WO-US16915.
XX
PR   18-JUN-1999; 99US-0140042.
XX

OM protein - protein search, using sw model

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

Run on: March 18, 2002, 09:55:33 ; Search sw model

Title: US-09-596-746A-34
Perfect score: 1809
Sequence: 1 MACNNSGKDGNTSANSADES.........KNLSKAAKEMLTNSVKELTS 371

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   Issued_Patents_AA:*
                 1:/cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
                 2:/cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
                 3:/cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
                 4:/cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
                 5:/cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
                 6:/cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Match Length	DB	ID	Description
1	702	38.8	466	4	US-08-235-836C-107	Sequence 107, App
2	701	38.8	588	4	US-08-235-836C-122	Sequence 122, App
3	692	38.3	466	4	PCT-US94-10261A-24	Sequence 110, App
4	658.5	36.4	209	4	US-09-196-293-15	Sequence 15, Appli
5	658.5	36.4	210	4	US-08-158-353-3	Sequence 3, Appli
6	653.5	36.1	210	4	US-08-235-836C-30	Sequence 30, Appli
7	652	36.0	207	4	US-08-235-836C-36	Sequence 36, Appli
8	651	36.0	210	4	US-08-209-603E-15	Sequence 15, Appli
9	597.5	33.0	212	1	US-08-158-353-2	Sequence 2, Appli
10	588	32.5	212	1	US-08-031-295-2	Sequence 2, Appli
11	588	32.5	212	1	US-07-903-580-2	Sequence 2, Appli
12	588	32.5	212	1	US-08-158-353-4	Sequence 4, Appli
13	556	30.7	209	4	US-08-235-836C-32	Sequence 32, Appli
14	554	30.6	212	4	US-09-196-293-11	Sequence 11, Appli
15	554	30.6	212	4	US-08-209-603E-11	Sequence 11, Appli
16	551	30.5	212	4	US-08-235-836C-34	Sequence 34, Appli
17	166	9.2	1713	3	US-08-600-982-24	Sequence 24, Appl
18	166.5	9.2	1713	5	PCT-US94-10261A-24	Sequence 24, Appl
19	166.5	9.1	296	1	US-08-320-161-11	Sequence 11, Appli
20	164.5	9.1	296	1	US-08-137-175A-3	Sequence 3, Appli
21	164.5	9.1	296	4	US-08-479-017-3	Sequence 3, Appl
22	164.5	9.1	296	4	US-08-455-829-11	Sequence 11, Appl
23	164.5	9.1	296	4	US-08-235-836C-22	Sequence 22, Appl
24	163.5	9.0	2482	1	US-08-328-254-6	Sequence 6, Appl
25	163.5	9.0	3248	1	US-08-353-700-1	Sequence 1, Appli
26	163.5	9.0	3248	5	PCT-US95-16216-1	Sequence 1, Appli
27	159.5	8.8	281	1	US-07-941-523-23	Sequence 23, Appl
28	159.5	8.8	380	4	US-08-235-836C-118	Sequence 118, App
29	159.5	8.8	393	4	US-08-235-836C-114	Sequence 114, App
30	159.5	8.8	441	4	US-08-235-836C-120	Sequence 120, App
31	159.5	8.8	454	4	US-08-235-836C-116	Sequence 116, App
32	152.5	8.4	1786	4	US-08-973-462-8	Sequence 8, Appli
33	151.5	8.4	3111	4	US-08-460-309-4	Sequence 4, Appli
34	151.5	8.4	3111	2	US-09-141-047-4	Sequence 4, Appli
35	143.5	7.9	688	3	US-09-141-047-8	Sequence 8, Appli
36	140.5	7.8	1027	4	US-08-446-137B-2	Sequence 2, Appl
37	139.5	7.7	1561	4	US-08-894-017-23	Sequence 23, Appl
38	137	7.6	1147	4	US-08-470-260-5	Sequence 5, Appli
39	137	7.6	1147	4	US-08-471-491-5	Sequence 5, Appli
40	137	7.6	1147	4	US-08-466-662-5	Sequence 5, Appli
41	137	7.6	1566	2	US-08-687-956A-23	Sequence 23, Appl
42	137	7.6	3289	2	US-08-477-451-2	Sequence 2, Appli
43	136	7.5	1588	5	PCT-US93-07261-11	Sequence 11, Appl
44	136	7.5	1663	5	PCT-US93-07261-16	Sequence 16, Appl
45	134	7.4	1196	1	US-08-144-121-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-08-235-836C-107
Sequence 107, Application US/08235836C
Patent No. 6248562

GENERAL INFORMATION:
APPLICANT: Dunn, John J.
APPLICANT: Luft, Benjamin J.
TITLE OF INVENTION: No. 6248562e1 Chimeric Proteins Comprising
TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor
NUMBER OF SEQUENCES: 144
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brookhaven National Laboratory
STREET:
CITY: Upton
STATE: NY
COUNTRY: USA
ZIP: 11973
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/235,836C
FILING DATE: 29-APR-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/148,191
FILING DATE: 01-11-93
ATTORNEY/AGENT INFORMATION:
NAME: Bogosian, Margaret C.
REGISTRATION NUMBER: 25,324
REFERENCE/DOCKET NUMBER: BNL93-28A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 282-7338
TELEFAX: (516) 282-3729
INFORMATION FOR SEQ ID NO: 107:
SEQUENCE CHARACTERISTICS:
LENGTH: 466 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-235-836C-107

Qy   19 ESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDELAKAIGKKIKNDGSLDNEANRNE 78

Query Match   38.8%; Score 702; DB 4; Length 466;
Best Local Similarity   47.5%; Pred. No. 2.2e-43;
Matches 173; Conservative 54; Mismatches 119; Indels 18; Gaps 6;

RESULT 2

```
Qy   100 EVFKEDGKTLVSKKVTSKDKSSTEEKFNEKGEVSKRIITRADGTRLEYTGIKSDGSGRAK 159
             | :|| |  |  :  :|  :|: |       | :|: :  |
Db    79 SLIAGAYTISTLITQKLSKL--NGSEGLKEKIAAAKKCSEEFSTKLKDNHAQLGIQGVTD 136

Qy   160 EVLKGYVLEGILTAEKTTLVVKEGTYVTLSKNISK----SGEVSVELNDTDSSAATKKTPA 215
             |  | :|: :  :  || :: : |       :  :   ||:| |  | :| |  | |
Db   137 ENAKKAILKANAAGK------DKGVEELEKLSGSLESLSKAAKEMLANSVKELTSPVV 188

Qy   216 WNDSTSTLTISADSKRTKDLVFLTDGTITVQQYNTAGISLEGSASEI--KNLSELKNALK 273
         ::| |  |  |  | | |    | :: | :|:  : |  |  :|:| |  || ||
Db   189 HGNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLTSIDEL-AKAI 247

Qy   274 MANNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAI 333
         |: ::   |      |  | :| |         :|  |     |  :| :: |: |  |
Db   248 GKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAEIEKAKKCSEEFTA 307

Qy   334 GKKIHQNNGLDTEYNHNGSLLAGRYAISTLIKOKLDGLKNEG-LKEKIDAARKCSETFTN 392
         |  | ::   | |  :|| :|:  :| || |  | ||   | ||:|::|| | |  ||
Db   308 KLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLSKAAKEMLTNSVK 367

Qy   393 KLKAKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 452
         |  ||||||||||| | | |:|||||| |||||  ||||||||: ||||||||||| |||
Db   368 ELTS 371

Qy   453 ELTS 456
Db       
```

US-08-235-836C-122

```
Sequence 122, Application US/08235836C
Patent No. 6248562
GENERAL INFORMATION:
    APPLICANT: Dunn, John J.
    APPLICANT: Luft, Benjamin J.
    TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising
    TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor
    NUMBER OF SEQUENCES: 144
    CORRESPONDENCE ADDRESS:
        ADDRESSEE: Brookhaven National Laboratory
        STREET:
        CITY: Upton
        STATE: NY
        COUNTRY: USA
        ZIP: 11973
    COMPUTER READABLE FORM:
        MEDIUM TYPE: Floppy disk
        COMPUTER: IBM PC compatible
        OPERATING SYSTEM: PC-DOS/MS-DOS
        SOFTWARE: PatentIn Release #1.0, Version #1.25
    CURRENT APPLICATION DATA:
        APPLICATION NUMBER: US/08/235,836C
        FILING DATE: 29-APR-1994
        CLASSIFICATION: 435
    PRIOR APPLICATION DATA:
        APPLICATION NUMBER: US 08/148,191
        FILING DATE: 01-11-93
    ATTORNEY/AGENT INFORMATION:
        NAME: Bogosian, Margaret C.
        REGISTRATION NUMBER: 25,324
        REFERENCE/DOCKET NUMBER: BNL93-28A
    TELECOMMUNICATION INFORMATION:
        TELEPHONE: (516) 282-7338
        TELEFAX: (516) 282-3729
INFORMATION FOR SEQ ID NO: 122:
    SEQUENCE CHARACTERISTICS:
        LENGTH: 588 amino acids
        TYPE: amino acid
        TOPOLOGY: linear
        MOLECULE TYPE: protein
US-08-235-836C-122
```

RESULT 3

```
Query Match
Best Local Similarity  38.8%;  Score 701;  DB 4;  Length 588;
Matches 181;  Conservative 56;  Mismatches 86;  Indels 68;  Gaps 15;
```

```
Qy     6 SGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDELAKAIGKKIK 65
         |: ||| |  |  :| |         :|  |     |  :| :: |: |  |  :||||:
Db   119 SSKVTKKQGSITEETLKANKLD--SKKLTRSNGTTLEYSQI---TDADNATKAV-ETLK 171

Qy    66 NDGSLDNEANRNESLLAGAYTI-----STLITQKLSK------LNGSEGLKEKIAAAKK 113
         |  |  |  |:| :|: |  ||      || ||||  |       |  |  :||||: |:
Db   172 NSIKLEG------SLVVGKTTVEIKEGTVTLKREIEKDGKVKVPLNDTAGSNKK---- 219

Qy   114 CSEEFSTKLKDNHAQLGIQGVTDENAKK---------AILKANAAGKDKGVEELEKL 161
         ||||||||||||||||||||||||| |:||           |: ::   |      |
Db   220 ------TGKWEDSTSTLITSA----DSKKTKDLVFLTDGTITVQQYNTAG-----TSLEGS 265

Qy   162 SGSLESLSKAAKEMLANSVKELTSPVVHGNNSGKDGNTSANSADESVKGPNLTEISKKIT 221
         :: :|: :  :  || :: : |       :|  |      |  | :| |
Db   266 ASEIKNLSE-----LKNALK--GHPM--GNNSGKDGNTSANSADESVKGPNLTEISKKIT 316

Qy   222 ESNAVVLAVKEVETLLTSIDEL-AKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITK 280
         :||||||||||||||||:||||  ||:|  |
Db   317 DSNAVLLAVKEVEALLSSIDEIAAKAIGKKIHQNNGLDTEYNHNGSLLAGRYAISTLIKQ 376

Qy   281 KISAIKDSGELKAEIEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKG 340
         | |  :|  |  |: :| | |  ||  |||||||||||||||| | | |:|||||| |||
Db   377 KLDGLKNEG-LKEKIDAARKCSETFTNKLKEKHTDLGKEGVTDADAKEAILKTNGTKTKG 435

Qy   341 ADELEKLFESVKNLSKAAKEMLTNSVKELTS 371
         |  ||||||||: ||||||||||| ||||||
Db   436 AEELGKLFESVEVLSKAAKEMLANSVKELTS 466
```

US-08-235-836C-110

```
Sequence 110, Application US/08235836C
Patent No. 6248562
GENERAL INFORMATION:
    APPLICANT: Dunn, John J.
    APPLICANT: Luft, Benjamin J.
    TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising
    TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor
    NUMBER OF SEQUENCES: 144
    CORRESPONDENCE ADDRESS:
        ADDRESSEE: Brookhaven National Laboratory
        STREET:
        CITY: Upton
        STATE: NY
        COUNTRY: USA
        ZIP: 11973
    COMPUTER READABLE FORM:
        MEDIUM TYPE: Floppy disk
        COMPUTER: IBM PC compatible
        OPERATING SYSTEM: PC-DOS/MS-DOS
        SOFTWARE: PatentIn Release #1.0, Version #1.25
    CURRENT APPLICATION DATA:
        APPLICATION NUMBER: US/08/235,836C
        FILING DATE: 29-APR-1994
        CLASSIFICATION: 435
    PRIOR APPLICATION DATA:
        APPLICATION NUMBER: US 08/148,191
        FILING DATE: 01-11-93
    ATTORNEY/AGENT INFORMATION:
        NAME: Bogosian, Margaret C.
        REGISTRATION NUMBER: 25,324
        REFERENCE/DOCKET NUMBER: BNL93-28A
    TELECOMMUNICATION INFORMATION:
        TELEPHONE: (516) 282-7338
        TELEFAX: (516) 282-3729
INFORMATION FOR SEQ ID NO: 110:
    SEQUENCE CHARACTERISTICS:
        LENGTH: 466 amino acids
        TYPE: amino acid
```

; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-235-836C-110

Query Match   38.3%;   Score 692; DB 4; Length 466;
Best Local Similarity   48.3%;   Pred. No. 1.1e-42;
Matches 185; Conservative 41; Mismatches 123; Indels 34; Gaps 11;

Qy   1   MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEL-AKA 59
            ::|||||||||||||||||||||||||||||||||||||||||||||||||||||: |||
Db   17   ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAKA 76

Qy   60   IGKKIKNDGSLDNEANRNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAARKCSEEFS 119
            ||||| ::|||:||||||||||||||||||||||||||||||||||||| |||||||: ||
Db   77   IGKKIHQNNGLDTEYNHNGSLLAGRYAISTLIKQKLDGLK-NEGLKEKIDAAKKCSETFT 135

Qy   120   TKLKDNHAQLGIQGVTDENAKKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANS 179
            ||:| |||:|| :|:||:||||||:|||::||||||||||:|||||||||||||||
Db   136   NKLKAKHTDLGKEGVTDADAKEAILKTNGT-KTKGAEELGKLFESVEVLSKAAKEMLANS 194

Qy   180   VKELTSPVVHGNNSGKDGNTSA------NSADESVKGPNLTEISKKITESNA--VVLAVKEV 233
            ||||||||| |:||:||:||| ||:|||||||||||||||||||| |||||||
Db   195   VKELTSPVVAESPKKPKQNVSSLDEKNSVSVDLPGEMKVLVSKEKNKDGKYDLIATVDKL 254

Qy   234   ETLLTS------IDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLI--SNLITKKISA 284
            ||||| |:|:|||||:|:||:||||||||||||||||||| |||||||||
Db   255   ELKGTSDKNNGSGVLEGVKADKSKVKLTIS-----DDLGQTTLEVFKEDGKTLVSKKVTS 309

Qy   285   IKDSGELKAEIEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADEL 344
            |:|:|:|:|:|:|: ||:| |:|||||||||||||||||:|||||||||:||||||:
Db   310   -KDKSSTEEKFNEKGEVSEKIITRADG--TRLEYTGIKSDGSGRAKEVLKGYVLEGTLTA 366

Qy   345   EKLFESV---------KNLSKAAK 359
            :|| || ||:|::|
Db   367   EKMTLVVKEGTVTLSKNISKSGE 389

RESULT 4
US-09-196-293-15
; Sequence 15, Application US/09196293
; Patent No. 6183755
; GENERAL INFORMATION:
; CURRENT APPLICATION NUMBER: US/09/196,293
; CURRENT FILING DATE: 1998-11-19
; EARLIER APPLICATION NUMBER: US 08/209,603
; EARLIER FILING DATE: 1994-03-10
; EARLIER APPLICATION NUMBER: US 07/862,535
; EARLIER FILING DATE: 1992-06-19
; APPLICANT: Preac-Mursic, Vera
; TITLE OF INVENTION: Active proteins from Borrelia
; APPLICANT: Wilske, Bettina
; APPLICANT: Soutschek, Erwin
; APPLICANT: Motz, Manfred
; APPLICANT: Fuchs, Renate
; TITLE OF INVENTION: burgdorferi
; FILE REFERENCE: 738.001US2
; GENERAL INFORMATION:
; TITLE OF INVENTION: Active proteins from Borrelia
; EARLIER APPLICATION NUMBER: WO PCT/EP90/02282
; EARLIER FILING DATE: 1990-12-21
; EARLIER APPLICATION NUMBER: DE P39 42 728.5
; EARLIER FILING DATE: 1989-12-22
; EARLIER APPLICATION NUMBER: DE P40 18 988.0
; EARLIER FILING DATE: 1990-06-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 209
; TYPE: PRT
; ORGANISM: Borrelia burgdorferi
US-09-196-293-15

Query Match   36.4%;   Score 658.5; DB 4; Length 209;

Best Local Similarity   76.7%;   Pred. No. 1.1e-40;
Matches 145; Conservative 11; Mismatches 30; Indels 3; Gaps 3;

Qy   1   MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEL-AKA 59
            ::|||||||||||||||||||||||||||||||||||||||||||||||||||::| |||
Db   17   ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAKA 76

Qy   60   IGKKIKNDGSLDNEANRNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAARKCSEEFS 119
            ||||| ::|||:||||||||||||||||||||||||||||||||||||| |||||||: ||
Db   77   IGKKIHQNNGLDTEYNHNGSLLAGRYAISTLIKQKLDGLK-NEGLKEKIDAAKKCSETFT 135

Qy   120   TKLKDNHAQLGIQGVTDENAKKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANS 179
            ||:| |||:|| :|:||:||||||:|||::||||||||||:|||||||||||||||
Db   136   NKLKAKHTDLGKEGVTDADAKEAILKTNGT-KTKGAEELGKLFESVEVLSKAAKEMLANS 194

Qy   180   VKELTSPVV 188
            |||||||||
Db   195   VKELTSPVV 203

RESULT 5
US-08-158-353-3
; Sequence 3, Application US/08158353
; Patent No. 5620862
; GENERAL INFORMATION:
; APPLICANT: Padula, Steven J.
; TITLE OF INVENTION: Methods for Diagnosing Early Lyme
; TITLE OF INVENTION: Disease
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: MA
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/158,353
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Carroll, Alice O.
; REGISTRATION NUMBER: 33,542
; REFERENCE/DOCKET NUMBER: UCT93-05
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 210 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-158-353-3

Query Match   36.4%;   Score 658.5; DB 1; Length 210;
Best Local Similarity   76.7%;   Pred. No. 1.1e-40;
Matches 145; Conservative 11; Mismatches 30; Indels 3; Gaps 3;

Qy   1   MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEL-AKA 59
            ::|||||||||||||||||||||||||||||||||||||||||||||||||||::| |||
Db   17   ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAKA 76

Qy   60   IGKKIKNDGSLDNEANRNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAARKCSEEFS 119
            ||||| ::|||:||||||||||||||||||||||||||||||||||||| |||||||: ||
Db   77   IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLK-NEGLKEKIDAAKKCSETFT 135

```
Qy  120  TKLKDNHAQLGIQGVTDENAKKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANS 179
         || ||| :|| :||||:||||| | || |||:||||||  |||||||||||||||||
Db  136  NKLKEKHTDLGKEGVTDADAKEAILKTNGT-KTKGAEELGKLFESVEVLSKAAKEMLANS 194

Qy  180  VKELTSPVV 188
         |||||||||
Db  195  VKELTSPVV 203
```

RESULT 6
US-08-235-836C-30
; Sequence 30, Application US/08235836C
; Patent No. 6248562
; GENERAL INFORMATION:
; APPLICANT:
; APPLICANT: Dunn, John J.
; APPLICANT: Luft, Benjamin J.
; TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor
; NUMBER OF SEQUENCES: 144
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Brookhaven National Laboratory
;   STREET:
;   CITY: Upton
;   STATE: NY
;   COUNTRY: USA
;   ZIP: 11973
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/235,836C
;   FILING DATE: 29-APR-1994
;   CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 08/148,191
;   FILING DATE: 01-11-93
; ATTORNEY/AGENT INFORMATION:
;   NAME: Bogosian, Margaret C.
;   REGISTRATION NUMBER: 25,324
;   REFERENCE/DOCKET NUMBER: BNL93-28A
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (516) 282-7338
;   TELEFAX: (516) 282-3729
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 210 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-235-836C-30

Query Match   36.1%; Score 653.5; DB 4; Length 210;
Best Local Similarity   76.2%; Pred. No. 2.5e-40;
Matches 144; Conservative 11; Mismatches 31; Indels   3; Gaps   3;

```
Qy    1  MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEL-AKA 59
         : ||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db   17  ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA 76

Qy   60  IGKKIKNDGSLDNEANRNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEEFS 119
         ||||| :|| |||| |||||||| ||||||||:||| ||:|||||||| |||||||| ||
Db   77  IGKKIHQNNGLDTEYNHNGSLLAGRYAISTLIKQKLDGLK-NEGLKEKIDAAKKCSEFFT 135

Qy  120  TKLKDNHAQLGIQGVTDENAKKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANS 179
         || ||| :|| :||||:||||| | || |||:||||||  |||||||||||||||||
Db  136  NKLKEKHTDLGKEGVTDADAKEAILKTNGT-KTKGAEELGKLFESVEVLSKAAKEMLANS 194

Qy  180  VKELTSPVV 188
         |||||||||
Db  195  VKELTSPVV 203
```

Db 195 VKELTSPVV 203

RESULT 7
US-08-235-836C-36
; Sequence 36, Application US/08235836C
; Patent No. 6248562
; GENERAL INFORMATION:
; APPLICANT:
; APPLICANT: Dunn, John J.
; APPLICANT: Luft, Benjamin J.
; TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor
; NUMBER OF SEQUENCES: 144
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Brookhaven National Laboratory
;   STREET:
;   CITY: Upton
;   STATE: NY
;   COUNTRY: USA
;   ZIP: 11973
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/235,836C
;   FILING DATE: 29-APR-1994
;   CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 08/148,191
;   FILING DATE: 01-11-93
; ATTORNEY/AGENT INFORMATION:
;   NAME: Bogosian, Margaret C.
;   REGISTRATION NUMBER: 25,324
;   REFERENCE/DOCKET NUMBER: BNL93-28A
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (516) 282-7338
;   TELEFAX: (516) 282-3729
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 207 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-235-836C-36

Query Match   36.0%; Score 652; DB 4; Length 207;
Best Local Similarity   74.5%; Pred. No. 3.1e-40;
Matches 140; Conservative 21; Mismatches 23; Indels   4; Gaps   3;

```
Qy    1  MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDELAKAI 60
         : |||||   |:|| |  |:||||||||:|||||||||| ||||||||||||||||||: |
Db   17  ISCNNSG--GDSASTNPDESAKGPNLTVISKKITDSNAFLLAVKEVEALLSSIDELSKAI 74

Qy   61  GKKIKNDGSLDNEANRNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEEFST 120
         |||||||| :||||||||||:|||||| ||:|||||| ||  |:| |||| |:|||:|||
Db   75  GKKIKNDGTLDNEANRNESLIAGAYEISKLITQKLSVLN-SEELKKIKEAKDCSEKFTT 133

Qy  121  KLKDNHAQLGIQGVTDENAKKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANSV 180
         ||||.||:||| .| :| ||||||||: ||  ||| |||| || |:|||||||||:||:| ||
Db  134  KLKDSHAELGIGSVQDDNAKKAILKTHGT-KDKGAKELEELFKSLESLSKAAQAALTNSV 192

Qy  181  KELTSPVV 188
         ||||:|||
Db  193  KELTNPVV 200
```

RESULT 8
US-08-209-603E-15
; Sequence 15, Application US/08209603E
; Patent No. 6248538

GENERAL INFORMATION:
APPLICANT: FUCHS, RENATE
APPLICANT: WILSKE, BETTINA
APPLICANT: PREAC-MURSIC, VERA
APPLICANT: MOTZ, MANFRED
APPLICANT: SOUTSCHEK, ERWIN
TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PROTEINS
TITLE OF INVENTION: FROM BORRELIA BURGDORFERI
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROOKS HAIDT HAFFNER & DELAHUNTY
STREET: 99 PARK AVENUE
CITY: NEW YORK
STATE: NY
COUNTRY: USA
ZIP: 10016
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" FLOPPY DISC
COMPUTER: AT&T - IBM COMPATIBLE
OPERATING SYSTEM: MS-DOS Version 6.2
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/209,603E
FILING DATE: 10-MAR-1994
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP90/02282
FILING DATE: 21-DEC-1990
APPLICATION NUMBER: US 07/862,535
FILING DATE: 19-JUN-1992
ATTORNEY/AGENT INFORMATION:
NAME: ROBINSON, WILLIAM R.
REGISTRATION NUMBER: 27,224
REFERENCE/DOCKET NUMBER: LKR-9217-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 697-3355
TELEFAX: (212) 557-5635
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 210
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
DESCRIPTION:
HYPOTHETICAL: N/A
ANTI-SENSE: N/A
FRAGMENT TYPE: N/A
ORIGINAL SOURCE:
ORGANISM: B. BURGDORFERI
IMMEDIATE SOURCE:
LIBRARY: DSM 5662
POSITION IN GENOME: N/A
FEATURE:
IDENTIFICATION METHOD: amino acid analysis
PUBLICATION INFORMATION: N/A
US-08-209-603E-15

Query Match   36.0%; Score 651.5; DB 4; Length 210;
Best Local Similarity   76.2%; Pred. No. 3.4e-40;
Matches 144; Conservative   11; Mismatches 31; Indels   3; Gaps   3;

```
Qy   1 MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEL-AKA  59
         :||||||||||||||||||||||||||||||||||||||||||:|||:|||:|
Db  17 ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVLEVEALLSSIDEIAKA  76

Qy  60 IGKKIKNDGSLDNEANRNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEEFS 119
        ||||  : ||||||||||||||||||
Db  77 IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLGK-NEGLKEKIDAAKKCSETT 135

Qy 120 TKLKDNHAQLGIQGVTDENAKKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANS 179
        |||:|   :||:|  :|||:||||
Db 136 NKLKEKHPDLGKEGVTDADAKEAILKTNGT-KTKGAEELGKLFESVEVLSKAAKEMLANS 194
```

RESULT 9
US-08-158-353-2
Sequence 2, Application US/08158353
Patent No. 5520862
GENERAL INFORMATION:
APPLICANT: Padula, Steven J.
TITLE OF INVENTION: Methods for Diagnosing Early Lyme
TITLE OF INVENTION: Disease
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Militia Drive
CITY: Lexington
STATE: MA
COUNTRY: USA
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/158,353
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Carroll, Alice O.
REGISTRATION NUMBER: 33,542
REFERENCE/DOCKET NUMBER: UCT93-05
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-861-6240
TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 212 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-158-353-2

Query Match   33.0%; Score 597.5; DB 1; Length 212;
Best Local Similarity   69.1%; Pred. No. 2.8e-36;
Matches 134; Conservative   20; Mismatches 35; Indels   5; Gaps   4;

```
Qy   1 MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDELA-KA  59
         :|||||||||||||||||||||||||||||||||||||||| |||:|||:|||:|: |
Db  17 ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVLEVEALLSSIDEVAKKA  76

Qy  60 IGKKIKNDGSLDNEANRNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEEFS 119
        ||  :  | |  |:  |  : ||||||||:|||||  :||: | ::|||||:  :|| |:
Db  77 IGNLIAQNG-LNAGANQNGSLLAGAYVISTLIAEKLDGLKNSEELKEKIEDAKKCNKAFT 135

Qy 120 TKLKDNHAQLGIQ--GVTDENAKKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLA 177
        .|||.:|| |||    .||:|||| :|||.|:| ||||::|||||| ::.:||||::.|:
Db 136 DKLKSSHAELGIANGAATDANAKAAILKTNGT-KDKGAQELEKLFESVKNLSKAAQETLN 194

Qy 178 NSVKELTSPVVHGN 191
        ||||||||||| |:
Db 195 NSVKELTSPVVAEN 208
```

Qy 180 VKELTSPVV 188
       |||||||||
Db 195 VKELTSPVV 203

RESULT 10
US-08-031-295-2
Sequence 2, Application US/08031295
Patent No. 5530103

GENERAL INFORMATION:
APPLICANT: LIVEY, Ian
APPLICANT: DORNER, Freidrich
TITLE OF INVENTION: METHOD AND COMPOSITION FOR THE
TITLE OF INVENTION: PREVENTION OF LYME DISEASE
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/031,295
FILING DATE: 19930312
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/903,580
FILING DATE: 25-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/824,161
FILING DATE: 22-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/727,245
FILING DATE: 11-JUL-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/142 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 212 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-031-295-2

Query Match   32.5%;   Score 588;   DB 1;   Length 212;
Best Local Similarity 68.9%;   Pred. No. 1.3e-35;
Matches 126;   Conservative 23;   Mismatches 32;   Indels 2;   Gaps 2;

```
Qy  191 NNSGKDGNT-SANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLTSIDELA-KAIG 248
         |||| |::: | | |||| |||||||||::||||||||||::||||||| ||||
Db   20 NNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETLVSSIDELATKAIG 79

Qy  249 KKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAEIEKAKKCSEEFTAK 308
         ||| :: | ||||||||:|||:|| |:|:|||||:|||||:||||||:|:||
Db   80 KKIQONNGLGANADKNGSLLAGAYAISTLITEKLKALKNSGELKAKIEDARKCSEDFTKK 139

Qy  309 LKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLSKAAKEMLTNSVKE 368
         |: | ||||||:|| | :| :||||||| |: ||| :|||| | : |||||
Db  140 LAAGHAQLGIDGATDNDSKEAILKINGTKTKGAEELVKLSESVASLSKAAQEASANSVKE 199

Qy  369 LTS 371
         |||
Db  200 LTS 202
```

RESULT 11
US-07-903-580-2
Sequence 2, Application US/07903580
Patent No. 6221363
GENERAL INFORMATION:

APPLICANT: LIVEY, Ian
APPLICANT: DORNER, Freidrich
TITLE OF INVENTION: METHOD AND COMPOSITION FOR THE
TITLE OF INVENTION: PREVENTION OF LYME DISEASE
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: Virginia
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/903,580
FILING DATE: 19920625
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/824,161
FILING DATE: 22-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/727,245
FILING DATE: 11-JUL-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/131 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 212 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-903-580-2

Query Match   32.5%;   Score 588;   DB 4;   Length 212;
Best Local Similarity 68.9%;   Pred. No. 1.3e-35;
Matches 126;   Conservative 23;   Mismatches 32;   Indels 2;   Gaps 2;

```
Qy  191 NNSGKDGNT-SANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLTSIDELA-KAIG 248
         |||| |::: | | |||| |||||||||::||||||||||::||||||| ||||
Db   20 NNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETLVSSIDELATKAIG 79

Qy  249 KKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAEIEKAKKCSEEFTAK 308
         ||| :: | ||||||||:|||:|| |:|:|||||:|||||:||||||:|:||
Db   80 KKIQONNGLGANADKNGSLLAGAYAISTLITEKLKALKNSGELKAKIEDARKCSEDFTKK 139

Qy  309 LKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLSKAAKEMLTNSVKE 368
         |: | ||||||:|| | :| :||||||| |: ||| :|||| | : |||||
Db  140 LAAGHAQLGIDGATDNDSKEAILKINGTKTKGAEELVKLSESVASLSKAAQEASANSVKE 199

Qy  369 LTS 371
         |||
Db  200 LTS 202
```

RESULT 12
US-08-158-353-4
Sequence 4, Application US/08158353
Patent No. 5620862
GENERAL INFORMATION:
APPLICANT: Padilla, Steven J.
TITLE OF INVENTION: Methods for Diagnosing Early Lyme
TITLE OF INVENTION: Disease

NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
  ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
  STREET: Two Militia Drive
  CITY: Lexington
  STATE: MA
  COUNTRY: USA
  ZIP: 02173
COMPUTER READABLE FORM:
  MEDIUM TYPE: Floppy disk
  COMPUTER: IBM PC compatible
  OPERATING SYSTEM: PC-DOS/MS-DOS
  SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/158,353
  FILING DATE:
  CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
  NAME: Carroll, Alice O.
  REGISTRATION NUMBER: 33,542
  REFERENCE/DOCKET NUMBER: UCT93-05
TELECOMMUNICATION INFORMATION:
  TELEPHONE: 617-861-6240
  TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 4:
  SEQUENCE CHARACTERISTICS:
  LENGTH: 212 amino acids
  TYPE: amino acid
  STRANDEDNESS: single
  TOPOLOGY: linear
  MOLECULE TYPE: protein
US-08-158-353-4

Query Match   30.7%;   Score 556;   DB 1;   Length 212;
Best Local Similarity   67.2%;   Pred. No. 2.8e-33;
Matches 123;   Conservative 19;   Mismatches 39;   Indels 2;   Gaps 2;

```
Qy  191  NNSGKDGNT-SANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLTSIDELA-KAIG  248
              ||||||| |:: |  ||||||||||||||:||||||:|||||||||||| ||||
Db   20  NNSGKGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETLVLSIDELAKKAIG   79

Qy  249  KKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAEIEKAKKCSEEFTAK  308
Db   80  OKIDNNNGLAALNNQNGSLLAGAYAISTLITEKLSKLKNLEELKTEIARAKKCSEEFTNK  139

Qy  309  LKGEHTDLGKEGVTDDNAKKAAILKTNNDKTKGADELEKLFESVKNLSRAAKEMLTNSVKE  368
Db  140  LKSGHADLGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGLLKAAQVALTNSVKE  199

Qy  369  LTS  371
Db  200  LTS  202
```

RESULT 13
US-08-235-836C-32
Sequence 32, Application US/08235836C
Patent No. 6248562
GENERAL INFORMATION:
  APPLICANT: Dunn, John J.
  APPLICANT: Luft, Benjamin J.
  TITLE OF INVENTION: No. 6248562e1 Chimeric Proteins Comprising
  TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor
  NUMBER OF SEQUENCES: 144
CORRESPONDENCE ADDRESS:
  ADDRESSEE: Brookhaven National Laboratory
  STREET:
  CITY: Upton
  STATE: NY
  COUNTRY: USA
  ZIP: 11973
COMPUTER READABLE FORM:

  MEDIUM TYPE: Floppy disk
  COMPUTER: IBM PC compatible
  OPERATING SYSTEM: PC-DOS/MS-DOS
  SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/235,836C
  FILING DATE: 29-APR-1994
  CLASSIFICATION: 435
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 08/148,191
  FILING DATE: 01-11-93
ATTORNEY/AGENT INFORMATION:
  NAME: Bogoslan, Margaret C.
  REGISTRATION NUMBER: 25,324
  REFERENCE/DOCKET NUMBER: BNL93-28A
TELECOMMUNICATION INFORMATION:
  TELEPHONE: (516) 282-3729
  TELEFAX: (516) 282-7338
INFORMATION FOR SEQ ID NO: 32:
  SEQUENCE CHARACTERISTICS:
  LENGTH: 209 amino acids
  TYPE: amino acid
  TOPOLOGY: linear
  MOLECULE TYPE: protein
US-08-235-836C-32

Query Match   30.6%;   Score 554;   DB 4;   Length 209;
Best Local Similarity   64.6%;   Pred. No. 3.8e-33;
Matches 122;   Conservative 27;   Mismatches 36;   Indels 4;   Gaps 3;

```
Qy    1  MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDELA-KA   59
         :||||| |:: ||  ||||||||||||||||||:||||||||||||||||||||| ||
Db   17  ISCNNSG--GDTASTNPDESAKGPNLTVISKKITDSNAFVLAVKEVEALLSSIDELANKA   74

Qy   60  IGKKIKNDGSLDNEANRNESLLAGAYTISTLITQKLSKLNGSEGLREKIAAAKKCSEEFS  119
Db   75  IGKVIHONGLNANAGQNGSLLAGAYAISTLITEKLSKLKNSEELNKKIEEAKNHSEAFT  134

Qy  120  TKLKDNHAQLGIQGVTDENAKKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANS  179
Db  135  NRLKGSHAQLGVAAATDDHAREAILKSNPT-KDKGAKALKDLSESVESLAKAAQEALANS  193

Qy  180  VKELTSPVV  188
Db  194  VKELTNPVV  202
```

RESULT 14
US-09-196-293-11
Sequence 11, Application US/09196293
Patent No. 6183755
GENERAL INFORMATION:
  APPLICANT: Fuchs, Renate
  APPLICANT: Motz, Manfred
  APPLICANT: Soutscheck, Erwin
  APPLICANT: Wilske, Bettina
  APPLICANT: Preac-Mursic, Vera
  TITLE OF INVENTION: Active proteins from Borrelia
  FILE REFERENCE: 738.001US2
  CURRENT APPLICATION NUMBER: US/09/196,293
  CURRENT FILING DATE: 1998-11-19
  TITLE OF INVENTION: burgdorferi
  EARLIER APPLICATION NUMBER: US 08/209,603
  EARLIER FILING DATE: 1994-03-10
  EARLIER APPLICATION NUMBER: US 07/862,535
  EARLIER FILING DATE: 1992-06-19
  EARLIER APPLICATION NUMBER: WO PCT/EP90/02282
  EARLIER FILING DATE: 1990-12-21
  EARLIER APPLICATION NUMBER: DE P39 42 728.5
  EARLIER FILING DATE: 1989-12-22
  EARLIER APPLICATION NUMBER: DE P40 18 988.0
  EARLIER FILING DATE: 1990-06-13

```
NUMBER OF SEQ ID NOS: 16
SOFTWARE: FastSEQ for Windows Version 4.0

SEQ ID NO 11
LENGTH: 212
TYPE: PRT
ORGANISM: Borrelia burgdorferi
US-09-196-293-11

Query Match    Best Local Similarity    30.6%;    Score 554;    DB 4;    Length 212;
Matches 124;    Conservative 17;    Mismatches 40;    Indels 2;    Gaps 2;

Qy  191  NNSGKDG-NTSANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLTSIDELA-KAIG  248
         |||||   | || ||||| ||||||||||||:||||:||||||||||: ||||||  |||
Db   20  NNSGKVGILTSTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETLVLSIDELAKKAIG  79

Qy  249  KKIKNDVSLDNEADHNGSLISGAYLISNLIITKKISAIKDSGELKAEIEKAKKCSEEFTAK  308
         :|| :  ::| | |||||| |||:||:|| || |  |:|  |:||:||:|||||||||| |
Db   80  QKIDNNNGLAALNNQNGSLLAGAYAISTLITEKLSKLKNLEELKTEIAKAKKCSEEFTNK  139

Qy  309  LKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLSKAAKEMLTNSVKE  368
         ||  | ||||  ||||:||||||||| :  |||:| |:||||| ||: |||:|||||||
Db  140  LKSGHADLGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGLLKAAQVALTNSVKE  199

Qy  369  LTS  371
         |||
Db  200  LTS  202

RESULT 15
US-08-209-603E-11
Sequence 11, Application US/08209603E
Patent No. 6248538
GENERAL INFORMATION:
     APPLICANT: FUCHS, RENATE
     APPLICANT: WILSKE, BETTINA
     APPLICANT: PREAC-MURSIC, VERA
     APPLICANT: MOTZ, MANFRED
     APPLICANT: SOUTSCHECK, ERWIN
     TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PROTEINS
     TITLE OF INVENTION: FROM BORRELIA BURGDORFERI
     NUMBER OF SEQUENCES: 15
     CORRESPONDENCE ADDRESS:
          ADDRESSEE: BROOKS HAIDT HAFFNER & DELAHUNTY
          STREET: 99 PARK AVENUE
          CITY: NEW YORK
          STATE: NY
          COUNTRY: USA
          ZIP: 10016
     COMPUTER READABLE FORM:
          MEDIUM TYPE: 3.5" FLOPPY DISC
          COMPUTER: AT&T - IBM COMPATIBLE
          OPERATING SYSTEM: MS-DOS Version 6.2
          SOFTWARE: ASCII
     CURRENT APPLICATION DATA:
          APPLICATION NUMBER: US/08/209,603E
          FILING DATE: 10-MAR-1994
          CLASSIFICATION: 436
     PRIOR APPLICATION DATA:
          APPLICATION NUMBER: PCT/EP90/02282
          FILING DATE: 21-DEC-1990
          APPLICATION NUMBER: US 07/862,535
          FILING DATE: 19-JUN-1992
     ATTORNEY/AGENT INFORMATION:
          NAME: ROBINSON, WILLIAM R.
          REGISTRATION NUMBER: 27,224
          REFERENCE/DOCKET NUMBER: LKR-9217-A
     TELECOMMUNICATION INFORMATION:
          TELEPHONE: (212) 697-3355
          TELEFAX: (212) 557-5635
INFORMATION FOR SEQ ID NO: 11:
     SEQUENCE CHARACTERISTICS:
          LENGTH: 212
          TYPE: AMINO ACID
          TOPOLOGY: LINEAR
     MOLECULE TYPE: PROTEIN
     DESCRIPTION: N/A
     HYPOTHETICAL: N/A
     ANTI-SENSE: N/A
     FRAGMENT TYPE: N/A
     ORIGINAL SOURCE:
          ORGANISM: B. BURGDORFERI
     IMMEDIATE SOURCE:
          LIBRARY: DSM 5662
     POSITION IN GENOME: N/A
     FEATURE:
          IDENTIFICATION METHOD: amino acid analysis
     PUBLICATION INFORMATION: N/A
     ANTI-SENSE: N/A
     FRAGMENT TYPE: INTERNAL
     ORIGINAL SOURCE:
US-08-209-603E-11

Query Match    Best Local Similarity    30.6%;    Score 554;    DB 4;    Length 212;
Matches 124;    Conservative 17;    Mismatches 40;    Indels 2;    Gaps 2;

Qy  191  NNSGKDG-NTSANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLTSIDELA-KAIG  248
         |||||   | || ||||| ||||||||||||:||||:||||||||||: ||||||  |||
Db   20  NNSGKVGILTSTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETLVLSIDELAKKAIG  79

Qy  249  KKIKNDVSLDNEADHNGSLISGAYLISNLIITKKISAIKDSGELKAEIEKAKKCSEEFTAK  308
         :|| :  ::| | |||||| |||:||:|| || |  |:|  |:||:||:|||||||||| |
Db   80  QKIDNNNGLAALNNQNGSLLAGAYAISTLITEKLSKLKNLEELKTEIAKAKKCSEEFTNK  139

Qy  309  LKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLSKAAKEMLTNSVKE  368
         ||  | ||||  ||||:||||||||| :  |||:| |:||||| ||: |||:|||||||
Db  140  LKSGHADLGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGLLKAAQVALTNSVKE  199

Qy  369  LTS  371
         |||
Db  200  LTS  202

Search completed: March 18, 2002, 09:55:34
Job time: 347 sec
```

THIS PAGE BLANK (USPTO)

OM protein - protein search, using sw model

  GenCore version 4.5
  Copyright (c) 1993 - 2000 Compugen Ltd.

Run on: March 18, 2002, 10:08:44 ; Search time 621.2 Seconds
             (without alignments)
             165.825 Million cell updates/sec

Title: US-09-596-746A-34
Perfect score: 1809
Sequence: 1 MACNNSGKDGNTSANSADES.........KNLSKAAKEMLTNSVKELTS 371

Scoring table: BLOSUM62
Searched: Gapop 10.0 , Gapext 0.5

      3148936 seqs, 27765703A residues

Total number of hits satisfying chosen parameters: 3148936

Post-processing: Minimum Match 0%
Maximum DB seq length: 200000000
Minimum DB seq length: 0
      Maximum Match 100%
      listing first 45 summaries

Database : Pending_Patents_AA_Main:*
   1: /cgn2_6/ptodata/2/paa/PCTUS_COMB.pep:*
   2: /cgn2_6/ptodata/2/paa/US06_COMB.pep:*
   3: /cgn2_6/ptodata/2/paa/US07_COMB.pep:*
   4: /cgn2_6/ptodata/2/paa/US80_COMB.pep:*
   5: /cgn2_6/ptodata/2/paa/US081_COMB.pep:*
   6: /cgn2_6/ptodata/2/paa/US082_COMB.pep:*
   7: /cgn2_6/ptodata/2/paa/US083_COMB.pep:*
   8: /cgn2_6/ptodata/2/paa/US084_COMB.pep:*
   9: /cgn2_6/ptodata/2/paa/US085_COMB.pep:*
  10: /cgn2_6/ptodata/2/paa/US086_COMB.pep:*
  11: /cgn2_6/ptodata/2/paa/US087_COMB.pep:*
  12: /cgn2_6/ptodata/2/paa/US088_COMB.pep:*
  13: /cgn2_6/ptodata/2/paa/US089_COMB.pep:*
  14: /cgn2_6/ptodata/2/paa/US090_COMB.pep:*
  15: /cgn2_6/ptodata/2/paa/US091_COMB.pep:*
  16: /cgn2_6/ptodata/2/paa/US092_COMB.pep:*
  17: /cgn2_6/ptodata/2/paa/US093_COMB.pep:*
  18: /cgn2_6/ptodata/2/paa/US094_COMB.pep:*
  19: /cgn2_6/ptodata/2/paa/US095_COMB.pep:*
  20: /cgn2_6/ptodata/2/paa/US096_COMB.pep:*
  21: /cgn2_6/ptodata/2/paa/US097_COMB.pep:*
  22: /cgn2_6/ptodata/2/paa/US098_COMB.pep:*
  23: /cgn2_6/ptodata/2/paa/US099_COMB.pep:*
  24: /cgn2_6/ptodata/2/paa/US60_COMB.pep:*

SUMMARIES

Result No.	Score	% Query Match	Length	DB	DB ID	Description
1	1809	100.0	371	19	US-09-596-746A-34	Sequence 34, Appl
2	1804	99.7	370	19	US-09-596-746A-34	Sequence 34, Appl
3	1801	99.6	394	19	US-09-596-746A-64	Sequence 64, Appl
4	1801	99.6	395	19	US-09-596-746A-64	Sequence 64, Appl
5	1564.5	86.5	370	19	US-09-596-746A-36	Sequence 36, Appl
6	1559.5	86.2	369	19	US-09-596-746-36	Sequence 36, Appl
7	1556.5	86.0	393	19	US-09-596-746-66	Sequence 66, Appl
8	1556.5	86.0	394	19	US-09-596-746A-66	Sequence 66, Appl
9	1534	84.8	375	19	US-09-596-746A-24	Sequence 24, Appl

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

ALIGNMENTS

RESULT 1
US-09-596-746A-34
Sequence 34, Application US/0956746A
GENERAL INFORMATION:
APPLICANT: Dattwyler, Raymond J.
APPLICANT: Seinost, Gerald
APPLICANT: Dykhuizen, Daniel
APPLICANT: Luft, Benjamin J.
APPLICANT: Maria J.C. Gomes-Solecki
TITLE OF INVENTION: Groups of Borrelia burgdorferi and
TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
FILE REFERENCE: 2631.1002-001
CURRENT APPLICATION NUMBER: US/09/596,746A
CURRENT FILING DATE: 2000-06-19
PRIOR APPLICATION NUMBER: US 60/140,042
PRIOR FILING DATE: 1999-06-18
NUMBER OF SEQ ID NOS: 84
SOFTWARE: FastSEQ for Windows Version 4.0

SEQ ID NO 34
LENGTH: 371
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: OspC Chimera
US-09-596-746A-34

Query Match   100.0%; Score 1809; DB 19; Length 371;
Best Local Similarity 100.0%; Pred. No. 1.9e-120;
Matches 371; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDELAKAI 60

10	1533.5	84.8	378	19	US-09-596-746-28	Sequence 28, Appl
11	1533.5	84.8	378	19	US-09-596-746A-28	Sequence 28, Appl
12	1529	84.5	374	19	US-09-596-746A-24	Sequence 24, Appl
13	1526	84.4	398	19	US-09-596-746A-60	Sequence 60, Appl
14	1526	84.4	398	19	US-09-596-746A-60	Sequence 60, Appl
15	1524.5	84.3	401	19	US-09-596-746A-80	Sequence 80, Appl
16	1524.5	84.3	402	19	US-09-596-746A-80	Sequence 80, Appl
17	1494	82.6	370	19	US-09-596-746A-40	Sequence 40, Appl
18	1492.5	82.5	370	19	US-09-596-746A-32	Sequence 32, Appl
19	1489	82.3	368	19	US-09-596-746-40	Sequence 40, Appl
20	1488.5	82.3	378	19	US-09-596-746A-54	Sequence 54, Appl
21	1488.5	82.3	369	19	US-09-596-746A-32	Sequence 32, Appl
22	1487.5	82.2	369	19	US-09-596-746A-54	Sequence 54, Appl
23	1486	82.1	378	19	US-09-596-746A-72	Sequence 72, Appl
24	1486	82.1	393	19	US-09-596-746A-72	Sequence 72, Appl
25	1484.5	82.1	393	19	US-09-596-746A-68	Sequence 68, Appl
26	1484.5	82.1	394	19	US-09-596-746A-68	Sequence 68, Appl
27	1480.5	81.8	401	19	US-09-596-746A-70	Sequence 70, Appl
28	1480.5	81.8	402	19	US-09-596-746A-70	Sequence 70, Appl
29	1289.5	71.3	374	19	US-09-596-746A-26	Sequence 26, Appl
30	1289	71.3	377	19	US-09-596-746A-30	Sequence 30, Appl
31	1289	71.3	377	19	US-09-596-746A-30	Sequence 30, Appl
32	1284.5	71.0	373	19	US-09-596-746A-26	Sequence 26, Appl
33	1281.5	70.8	397	19	US-09-596-746A-62	Sequence 62, Appl
34	1281.5	70.8	398	19	US-09-596-746A-62	Sequence 62, Appl
35	1281.5	70.8	400	19	US-09-596-746A-82	Sequence 82, Appl
36	1280	70.8	401	19	US-09-596-746A-82	Sequence 82, Appl
37	1249.5	69.1	368	19	US-09-596-746A-42	Sequence 42, Appl
38	1244.5	68.8	367	19	US-09-596-746A-42	Sequence 42, Appl
39	1241.5	68.6	391	19	US-09-596-746A-74	Sequence 74, Appl
40	1241.5	68.6	392	19	US-09-596-746A-74	Sequence 74, Appl
41	1230	68.0	385	19	US-09-596-746A-56	Sequence 56, Appl
42	1225	67.7	384	19	US-09-596-746A-56	Sequence 56, Appl
43	1222	67.6	408	19	US-09-596-746A-78	Sequence 78, Appl
44	1222	67.6	409	19	US-09-596-746A-78	Sequence 78, Appl
45	1199	66.3	560	19	US-09-596-746A-52	Sequence 52, Appl

Left column (RESULT 2)

RESULT 2
US-09-596-746-34
Sequence 34, Application US/0959674b
GENERAL INFORMATION:
APPLICANT: Dattwyler,
APPLICANT: Seinost, Gerald
APPLICANT: Dykhuizen, Raymond J.
APPLICANT: Luft, Benjamin J.
APPLICANT: Maria J.C. Gomes-Solecki
TITLE OF INVENTION: Groups of Borrelia burgdorferi and Borrelia afzelii That Cause Lyme Disease in Humans
FILE REFERENCE: 2631.1002-001
CURRENT APPLICATION NUMBER: US/09/596,746
PRIOR APPLICATION NUMBER: US 60/140,042
PRIOR FILING DATE: 2000-06-16
NUMBER OF SEQ ID NOS: 84
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO: 34
LENGTH: 370
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: OspC Chimera
US-09-596-746-34

Query Match 99.7%; Score 1804; DB 19; Length 370;
Best Local Similarity 100.0%; Pred. No. 4.2e-120;
Matches 370; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    1 MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDELAKAI  60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   61 GKKIKNDGSLDNEANRNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEEFST 120
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 KKIKNDGSLDNEANRNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEEFST  120
Qy  121 KLKDNHAQLGIQGVTDENAKKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANSV 180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 KLKDNHAQLGIQGVTDENAKKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANSV 180
Qy  181 KELTSPVVHGNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLTSI 240
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 KELTSPVVHGNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLTSI 240
Qy  241 DELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAEIEKAKK 300
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 DELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAEIEKAKK 300
Qy  301 CSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLSKAAKE 360
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 CSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLSKAAKE 360
Qy  361 MLTNSVKELTS 371
        |||||||||||
Db  361 MLTNSVKELTS 371
```

Right column (RESULT 3)

RESULT 3
US-09-596-746-64
Sequence 64, Application US/0959674b
GENERAL INFORMATION:
APPLICANT: Dattwyler,
APPLICANT: Seinost, Gerald
APPLICANT: Dykhuizen, Raymond J.
APPLICANT: Luft, Benjamin J.
APPLICANT: Maria J.C. Gomes-Solecki
TITLE OF INVENTION: Groups of Borrelia burgdorferi and Borrelia afzelii That Cause Lyme Disease in Humans
FILE REFERENCE: 2631.1002-001
CURRENT APPLICATION NUMBER: US/09/596,746
PRIOR APPLICATION NUMBER: US 60/140,042
PRIOR FILING DATE: 2000-06-16
NUMBER OF SEQ ID NOS: 84
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO: 64
LENGTH: 394
TYPE: PRT
ORGANISM: OspC Chimera
US-09-596-746-64

Query Match 99.6%; Score 1801; DB 19; Length 394;
Best Local Similarity 99.7%; Pred. No. 7.6e-120;
Matches 369; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy    2 ACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDELAKAIG  61
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   25 SCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDELAKAIG  84
Qy   62 KKIKNDGSLDNEANRNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEEFSTK 121
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   85 KKIKNDGSLDNEANRNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEEFSTK 144
Qy  122 LKDNHAQLGIQGVTDENAKKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANSVK 181
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  145 LKDNHAQLGIQGVTDENAKKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANSVK 204
Qy  182 ELTSPVVHGNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLTSID 241
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  205 ELTSPVVHGNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLTSID 264
Qy  242 ELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAEIEKAKKC 301
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  265 ELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAEIEKAKKC 324
Qy  302 SEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLSKAAKEM 361
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  325 SEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLSKAAKEM 384
Qy  362 LTNSVKELTS 371
        ||||||||||
Db  385 LTNSVKELTS 394
```

RESULT 4
US-09-596-746-64
Sequence 64, Application US/0959674A
GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seinost, Gerald
; APPLICANT: Dykhuizen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746A
; CURRENT FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 64
; TYPE: PRT
; ORGANISM: OspC Chimera
; LENGTH: 395

US-09-596-746A-64

Query Match   99.6%;   Score 1801;   DB 19;   Length 395;
Best Local Similarity   99.7%;   Pred. No. 7.6e-120;
Matches 369; Conservative   1; Mismatches   0;   Indels   0;   Gaps   0;

Qy   2   ACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDELAKAIG   61
        :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   26   SCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDELAKAIG   85

Qy   62   KKIKNDGSLDNEANRNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCCSEEFSTK   121
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   86   KKIKNDGSLDNEANRNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCCSEEFSTK   145

Qy   122   LKDNHAQLGIGQGVTDENAKKAILKANAAGKDRGVEELEKLSGSSLESLSKAAKEMLANSVK   181
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   146   LKDNHAQLGIGQGVTDENAKKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANSVK   205

Qy   182   ELTSPVVHGNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLTSID   241
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   206   ELTSPVVHGNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLTSID   265

Qy   242   ELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAEIEKAKKC   301
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   266   ELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAEIEKAKKC   325

Qy   302   SEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLSKAAKEM   361
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   326   SEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLSKAAKEM   385

Qy   362   LTNSVKELTS   371
        ||||||||||
Db   386   LTNSVKELTS   395

RESULT 5
US-09-596-746A-36
Sequence 36, Application US/0959674A
GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seinost, Gerald
; APPLICANT: Dykhuizen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: 2000-06-19
; PRIOR APPLICATION NUMBER: 2000-06-19
; PRIOR FILING DATE: 2000-06-19

; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 36
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: OspC Chimera
; SEQ ID NO 36
; LENGTH: 370
US-09-596-746A-36

Query Match   86.5%;   Score 1564.5;   DB 19;   Length 370;
Best Local Similarity   87.6%;   Pred. No. 5e-103;
Matches 324; Conservative   20; Mismatches   25;   Indels   1;   Gaps   1;

Qy   1   MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDELAKAI   60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1   MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDELAKAI   60

Qy   61   GKKIKNDGSLDNEANRNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCCSEEFST   120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61   GKKIKNDGSLDNEANRNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCCSEEFST   120

Qy   121   KLKDNHAQLGIGQGVTDENAKKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANSV   180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121   KLKDNHAQLGIGQGVTDENAKKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANSV   180

Qy   181   KELTSPVVHGNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLTSI   240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181   KELTSPVVHGNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASI   240

Qy   241   DELA-KAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAEIEKAK   299
        |||| ||||||||||||||||:||||||||||||:|:||:||||||||||||||||||||
Db   241   DELATRAIGKKIQQNGGLAVEAGHNGTLLAGAYTISKLIITQKLDGLKNSEKIKEKIENAK   300

Qy   300   KCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLSKAAK   359
        ||||:|||:||||||||||||||||:|||:|:|:|:|||||||||||||||||||:||||
Db   301   KCSEDTTKKLEGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLAKAAK   360

Qy   360   EMLTNSVKEL   369
        ||| ||||||
Db   361   EMLANSVKEL   370

RESULT 6
US-09-596-746-36
Sequence 36, Application US/0959674
GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seinost, Gerald
; APPLICANT: Dykhuizen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746
; CURRENT FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 36
; LENGTH: 369
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: OspC Chimera
US-09-596-746-36

Query Match   86.2%;   Score 1559.5;   DB 19;   Length 369;

RESULT 7

```
Best Local Similarity 87.5%;   Pred. No. 1.1e-102;
Matches 323;  Conservative 20;  Mismatches 25;  Indels 1;  Gaps 1;

Qy    2 ACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDELAKAIG  61
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 ACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDELAKAIG  60

Qy   62 KKIKNDGSLDNEANRNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEEFSTK 121
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 KKIKNDGSLDNEANRNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEEFSTK 120

Qy  122 LKDNHAQLGIQGVTDENAKKAIILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANSVK 181
        ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db  121 LKDNHAQLGIQGVTDENAKKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANSVK 180

Qy  182 ELTSPVVHGNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLTSID 241
        ||||||||||||||||||||||||||||||||||||||||||||||||||:||| |||
Db  181 ELTSPVVHGNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASID 240

Qy  242 ELA-KAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAEIEKAKK 300
        ||| ||||||| : | : ||:::|||:| :|| ::|| :|| :| ||:|| ||| |||
Db  241 ELATKAIGKKIQQNGGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKEKIENAKK 300

Qy  301 CSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLSKAAKE 360
        |||:|| :| |||| ||| |:|| ||||| :|| ||:|| ||||||| :|:||:||||
Db  301 CSEDFTKKLEGEHAQLGIENVTDENAKKAILTDAAKDKGAAELEKLFKAVENLAKAAKE 360

Qy  361 MLTNSVKEL 369
        || |||||
Db  361 MLANSVKEL 369
```

```
RESULT 7
US-09-596-746-66
; Sequence 66, Application US/0959746
; GENERAL INFORMATION:
;   APPLICANT: Dattwyler, Raymond J.
;   APPLICANT: Seinost, Gerald
;   APPLICANT: Dykhuizen, Danial
;   APPLICANT: Luft, Benjamin J.
;   APPLICANT: Maria J.C. Gomes-Solecki
;   TITLE OF INVENTION: Groups of Borrelia burgdorferi and
;   TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
;   FILE REFERENCE: 2631.1002-001
;   CURRENT APPLICATION NUMBER: US/09/596,746
;   CURRENT FILING DATE: 2000-06-16
;   PRIOR APPLICATION NUMBER: US 60/140,042
;   PRIOR FILING DATE: 1999-06-18
;   NUMBER OF SEQ ID NOS: 84
;   SOFTWARE: FastSEQ for Windows Version 4.0
;   SEQ ID NO 66
;   LENGTH: 393
;   TYPE: PRT
;   ORGANISM: OspC Chimera
US-09-596-746-66

Query Match        86.0%;  Score 1556.5;  DB 19;  Length 393;
Best Local Similarity 87.3%;  Pred. No. 2e-102;
Matches 322;  Conservative 21;  Mismatches 25;  Indels 1;  Gaps 1;

Qy    2 ACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDELAKAIG  61
        :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   25 SCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDELAKAIG  84

Qy   62 KKIKNDGSLDNEANRNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEEFSTK 121
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   85 KKIKNDGSLDNEANRNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEEFSTK 144

Qy  122 LKDNHAQLGIQGVTDENAKKAIILKANANAGKDKGVEELEKLSGSLESLSKRAAKEMLANSVK 181
        ||||||||||||||||||||| |||||:| ||||||||||||||||||||| |||||||||
Db  145 LKDNHAQLGIQGVTDENAKKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANSVK 204
```

RESULT 8

```
RESULT 8
US-09-596-746a-66
; Sequence 66, Application US/0959746A
; GENERAL INFORMATION:
;   APPLICANT: Dattwyler, Raymond J.
;   APPLICANT: Seinost, Gerald
;   APPLICANT: Dykhuizen, Danial
;   APPLICANT: Luft, Benjamin J.
;   APPLICANT: Maria J.C. Gomes-Solecki
;   TITLE OF INVENTION: Groups of Borrelia burgdorferi and
;   TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
;   FILE REFERENCE: 2631.1002-001
;   CURRENT APPLICATION NUMBER: US/09/596,746A
;   CURRENT FILING DATE: 2000-06-19
;   PRIOR APPLICATION NUMBER: US 60/140,042
;   PRIOR FILING DATE: 1999-06-18
;   NUMBER OF SEQ ID NOS: 84
;   SOFTWARE: FastSEQ for Windows Version 4.0
;   SEQ ID NO 66
;   LENGTH: 394
;   TYPE: PRT
;   ORGANISM: OspC Chimera
US-09-596-746a-66

Query Match        86.0%;  Score 1556.5;  DB 19;  Length 394;
Best Local Similarity 87.3%;  Pred. No. 2e-102;
Matches 322;  Conservative 21;  Mismatches 25;  Indels 1;  Gaps 1;

Qy    2 ACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDELAKAIG  61
        :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   26 SCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDELAKAIG  85

Qy   62 KKIKNDGSLDNEANRNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEEFSTK 121
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   86 KKIKNDGSLDNEANRNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEEFSTK 145

Qy  122 LKDNHAQLGIQGVTDENAKKAIILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANSVK 181
        ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db  146 LKDNHAQLGIQGVTDENAKKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANSVK 205

Qy  182 ELTSPVVHGNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLTSID 241
        ||||||||||||||||||||||||||||||||||||||||||||||||||:||| |||
Db  206 ELTSPVVHGNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASID 264

Qy  242 ELA-KAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAEIEKAKK 300
        ||| ||||||| : | : ||:::|||:| :|| ::|| :|| :| ||:|| ||| |||
Db  265 ELATKAIGKKIQQNGGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKEKIENAKK 325

Qy  301 CSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLSKAAKE 360
        |||:|| :| |||| ||| |:|| ||||| :|| ||:|| ||||||| :|:||:||||
Db  326 CSEDFTKKLEGEHAQLGIENVTDENAKKAILTDAAKDKGAAELEKLFKAVENLAKAAKE 385

Qy  361 MLTNSVKEL 369
        || |||||
Db  386 MLANSVKEL 394
```

RESULT 9
US-09-596-746A-24
Sequence 24, Application US/0959746A
GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seinost, Gerald
; APPLICANT: Dykhuizen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746A
; CURRENT FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: OspC Chimera
US-09-596-746A-24

Query Match   84.8%; Score 1534; DB 19; Length 375;
Best Local Similarity   87.0%; Pred. No. 7.5e-101;
Matches 328; Conservative   9; Mismatches 32; Indels   8; Gaps   4;

Qy   1 MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEL-AKA 59
          |||||||||||||||||||||||||||||||||||||||||||||||||| ||||: |||
Db   1 MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA 60

Qy   60 IGKKIKNDGSLDNEANRNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEEFS 119
          |||||||||||| |||||:|||||||| ||: |||||||||| ::||||||:|||||| |:
Db   61 IGKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLK-NEGLKEKIDAAKKCSETFT 119

Qy   120 TKLKDNHAQLGIQGYVTDENAKKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANS 179
          ||: || ||:|| |||||:||||||||| |||| ||||||||||||||||||||||
Db   120 NKLKEKHTDLGKEGVTDADAKEAILKYNGT-KTKGAEELGKLFESVEVLSKAAKEMLANS 178

Qy   180 VKELTSPVVHG------NNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVE 234
          ||||:|| | |||||||||||||||||||||||||||||||||||||||||||
Db   179 VKELTSPVVAESPAMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVE 238

Qy   235 TLLTSIDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAE 294
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   239 TLLTSIDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAE 298

Qy   295 IEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNL 354
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   299 IEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNL 358

Qy   355 SKAAKEMLTNSVKELTS 371
          |||||||||||||||||
Db   359 SKAAKEMLTNSVKELTS 375

RESULT 10
US-09-596-746A-28
Sequence 28, Application US/0959746A
GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seinost, Gerald
; APPLICANT: Dykhuizen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans

FILE REFERENCE: 2631.1002-001
CURRENT APPLICATION NUMBER: US/09/596,746A
CURRENT FILING DATE: 2000-06-16
PRIOR APPLICATION NUMBER: US 60/140,042
PRIOR FILING DATE: 1999-06-18
NUMBER OF SEQ ID NOS: 84
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 28
LENGTH: 378
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: OspC Chimera
US-09-596-746A-28

Query Match   84.8%; Score 1533.5; DB 19; Length 378;
Best Local Similarity   86.3%; Pred. No. 8.3e-101;
Matches 328; Conservative 10; Mismatches 31; Indels 11; Gaps   4;

Qy   1 MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEL-AKA 59
          |||||||||||||||||||||||||||||||||||||||||||||||||| ||||: |||
Db   1 MACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKA 60

Qy   60 IGKKIKNDGSLDNEANRNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEEFS 119
          |||||||||||| |||||:|||||||| ||: |||||||||| ::||||||:|||||| |:
Db   61 IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLK-NEGLKEKIDAAKKCSETFT 119

Qy   120 TKLKDNHAQLGIQGYVTDENAKKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANS 179
          ||: || ||:|| |||||:||||||||| |||| ||||||||||||||||||||||
Db   120 NKLKEKHTDLGKEGVTDADAKEAILKANGT-KTKGAEELGKLFESVEVLSKAAKEMLANS 178

Qy   180 VKELTSPVVHG-------NNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVK 231
          ||||:|| | |||||||||||||||||||||||||||||||||||||||||
Db   179 VKELTSPVVAESPKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVK 238

Qy   232 EVETLLTSIDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGEL 291
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   239 EVETLLTSIDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGEL 298

Qy   292 KAEIEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESV 351
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   299 KAEIEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESV 358

Qy   352 KNLSKAAKEMLTNSVKELTS 371
          ||||||||||||||||||||
Db   359 KNLSKAAKEMLTNSVKELTS 378

RESULT 11
US-09-596-746A-28
Sequence 28, Application US/0959746A
GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seinost, Gerald
; APPLICANT: Dykhuizen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746A
; CURRENT FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 28
; LENGTH: 378
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: OspC Chimera

US-09-596-746A-28

```
Query Match       84.8%;  Score 1533.5;  DB 19;  Length 378;
Best Local Similarity  86.3%;  Pred. No. 8.3e-101;
Matches 328;  Conservative 10;  Mismatches 31;  Indels 11;  Gaps 4;

Qy    1 MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEL-AKA   59
        ||||:||||||||||||||||||||||||: ||||||||||||||||||||||||: ||
Db    1 MACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEI-AKA   60

Qy   60 IGKKIKNDGSLDNEANRNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEEFS  119
        ||||| | |: |::|: |||||| |||||| ||:||  : |||||||| ||:|||||: |
Db   61 IGKKIHQNNGLDTENVNHNGSLLAGAYAISTLIKQKLDGLK-NEGLKEKIDAAKKCSETFT 119

Qy  120 TKLKDNHAQLGIQGVTDENAKKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANS  179
        .||| ||: ||| ||||| |||||||||  | |||||| ||| ||: |||||||||||||
Db  120 NKLKEKHTDLGKEGVTDADAKEAILKANGT-KTKGAEELGKLFESVEVLSKAAKEMLANS  178

Qy  180 VKELTSPVVHG--------NNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVK  231
        |||||||||           |||||||||||||||||||||||||||||||||||||||
Db  179 VKELTSPVAESPKKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVK   238

Qy  232 EVETLLTSIDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGEL  291
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  239 EVETLLTSIDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGEL  298

Qy  292 KAEIERAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESV  351
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  299 KAEIERAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESV  358

Qy  352 KNLSKAAKEMLTNSVKELTS  371
        ||||||||||||||||||||
Db  359 KNLSKAAKEMLTNSVKELTS  373
```

RESULT 12
US-09-596-746-24
Sequence 24, Application US/0959746
GENERAL INFORMATION:
APPLICANT: Dattwyler, Raymond J.
APPLICANT: Seinost, Gerald
APPLICANT: Dykhuizen, Danial
APPLICANT: Luft, Benjamin J.
APPLICANT: Maria J.C. Gomes-Solecki
TITLE OF INVENTION: Groups of Borrelia burgdorferi and
TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
FILE REFERENCE: 2631.1002-001
CURRENT APPLICATION NUMBER: US/09/596,746
CURRENT FILING DATE: 2000-06-16
PRIOR APPLICATION NUMBER: US 60/140,042
PRIOR FILING DATE: 1999-06-18
NUMBER OF SEQ ID NOS: 84
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 24
LENGTH: 374
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: OspC Chimera
US-09-596-746-24

```
Query Match       84.5%;  Score 1529;  DB 19;  Length 374;
Best Local Similarity  87.0%;  Pred. No. 1.7e-100;
Matches 327;  Conservative 9;  Mismatches 32;  Indels 8;  Gaps 4;

Qy    2 ACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEL-AKAI  60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||:  |||
Db    1 ACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAI  60

Qy   61 GKKIKNDGSLDNEANRNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEEFST 120
        |||| | |: |::|: |||||| |||||| ||:||  : |||||||| ||:|||||: |:
Db   61 GKKIHQNNGLDTEVNHNGSLLAGAYAISTLIKQKLDGLK-NEGLKEKIDAAKKCSETFTN 143
```

US-09-596-746-60

```
RESULT 13
US-09-596-746-60
Sequence 60, Application US/0959746
GENERAL INFORMATION:
APPLICANT: Dattwyler, Raymond J.
APPLICANT: Seinost, Gerald
APPLICANT: Dykhuizen, Danial
APPLICANT: Luft, Benjamin J.
APPLICANT: Maria J.C. Gomes-Solecki
TITLE OF INVENTION: Groups of Borrelia burgdorferi and
TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
FILE REFERENCE: 2631.1002-001
CURRENT APPLICATION NUMBER: US/09/596,746
CURRENT FILING DATE: 2000-06-16
PRIOR APPLICATION NUMBER: US 60/140,042
PRIOR FILING DATE: 1999-06-18
NUMBER OF SEQ ID NOS: 84
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 60
LENGTH: 398
TYPE: PRT
ORGANISM: ospC Chimera
US-09-596-746-60

Query Match       84.4%;  Score 1526;  DB 19;  Length 398;
Best Local Similarity  86.7%;  Pred. No. 3.1e-100;
Matches 326;  Conservative 10;  Mismatches 32;  Indels 8;  Gaps 4;

Db   61 GKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLK-NEGLKEKIDAAKKCSETFTN 119

Qy  121 KLKDNHAQLGIQGVTDENAKKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANSV 180
        |||| ||: ||| ||||| |||||||:||  | |||||| ||| ||: |||||||||||||
Db  120 KLKEKHTDLGKEGVTDADAKEAILKTNGT-KTKGAEELGKLFESVEVLSKAAKEMLANSV 178

Qy  181 KELTSPVVHG-----NNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVET 235
        |||||||||        |||||||||||||||||||||||||||||||||||||||||||
Db  120 KELTSPVVAESPANVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVET 178

Qy  236 LLTSIDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAEI 295
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  203 LLTSIDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAEI 262

Qy  296 ERAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLS 355
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  263 ERAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLS 322

Qy  356 KAAKEMLTNSVKELTS 371
        ||||||||||||||||
Db  359 KAAKEMLTNSVKELTS 374
```

```
Query Match       84.8%;  Score 1529;  DB 19;  Length 398;
Best Local Similarity  86.7%;  Pred. No.
Matches 326;  Conservative 10;  Mismatches 32;  Indels 8;  Gaps 4;

Qy    2 ACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEL-AKAI  60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||:  |||
Db   25 SCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAI  84

Qy   61 GKKIKNDGSLDNEANRNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEEFST 120
        |||| | |: |::|: |||||| |||||| ||:||  : |||||||| ||:|||||: |:
Db   85 GKKIHQNNGLDTEVNHNGSLLAGAYAISTLIKQKLDGLK-NEGLKEKIDAAKKCSETFTN 143
```

```
US-09-596-746-60

Query Match       84.4%;
Best Local Similarity  86.7%;  Pred. No.
Matches 326;  Conservative 10;
```

Db   323 EKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLS 382
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   356 KAAKEMLTNSVKELTS 371
           |||||||||||||||
Db   383 KAAKEMLTNSVKELTS 398

RESULT 14
US-09-596-746A-60
;   Sequence 60, Application US/0959674A
;   GENERAL INFORMATION:
;   APPLICANT: Dattwyler, Raymond J.
;   APPLICANT: Selnost, Gerald
;   APPLICANT: Dykhuizen, Danial
;   APPLICANT: Luft, Benjamin J.
;   APPLICANT: Maria J.C. Gomes-Solecki
;   TITLE OF INVENTION: Groups of Borrelia burgdorferi and
;   TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
;   FILE REFERENCE: 2631.1002-001
;   CURRENT APPLICATION NUMBER: US/09/596,746A
;   CURRENT FILING DATE: 2000-06-19
;   PRIOR APPLICATION NUMBER: US 60/140,042
;   PRIOR FILING DATE: 1999-06-18
;   NUMBER OF SEQ ID NOS: 84
;   SOFTWARE: FastSEQ for Windows Version 4.0
;   SEQ ID NO 60
;   LENGTH: 399
;   TYPE: PRT
;   ORGANISM: OspC Chimera
US-09-596-746A-60

Query Match   84.4%;   Score 1526; DB 19; Length 399;
Best Local Similarity   86.7%;   Pred. No. 3.1e-100;
Matches 326; Conservative 10; Mismatches 32; Indels 8; Gaps 4;

Qy   2 ACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEL-AKAI 60
           :|||||||||||||||||||||||||||||||||||||||||||||||||||: |||
Db   26 SCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAI 85

Qy   61 GKKIKNDGSLDNEANRNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEEFST 120
           ||| : ||| | |||||||||||||||| |||:| ||||| | |||||||||||||||: :
Db   86 GKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLK-NEGLKEKIDAAKKCSETFTN 144

Qy   121 KLKDNHAQLGIQGVTDENAKKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANSV 180
           ||| : || | :|||| |||||||||| || |||||||||||||||||||||||||||||
Db   145 KLKEKHTDLGKEGVTDADAKEAILKTNGT-KTRGAEELGKLFESVEVLSKAAKEMLANSV 203

Qy   181 KELTSPVHG------NNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVET 235
           ||||||| |||||||||||||||||||||||||||||||||||||||||||
Db   204 KELTSPVVAESPAMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVET 263

Qy   236 LLTSIDELAKAIGKKIRNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAEI 295
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   264 LLTSIDELAKAIGKKIRNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAEI 323

Qy   296 EKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLS 355
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   324 EKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLS 383

Qy   356 KAAKEMLTNSVKELTS 371
           ||||||||||||||||
Db   384 KAAKEMLTNSVKELTS 399

RESULT 15
US-09-596-746-80
;   Sequence 80, Application US/0959746
;   GENERAL INFORMATION:
;   APPLICANT: Dattwyler, Raymond J.
;   APPLICANT: Selnost, Gerald
;   APPLICANT: Dykhuizen, Danial
;   APPLICANT: Luft, Benjamin J.
;   APPLICANT: Maria J.C. Gomes-Solecki
;   TITLE OF INVENTION: Groups of Borrelia burgdorferi and
;   TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
;   FILE REFERENCE: 2631.1002-001
;   CURRENT APPLICATION NUMBER: US/09/596,746
;   CURRENT FILING DATE: 2000-06-16
;   PRIOR APPLICATION NUMBER: US 60/140,042
;   PRIOR FILING DATE: 1999-06-18
;   NUMBER OF SEQ ID NOS: 84
;   SOFTWARE: FastSEQ for Windows Version 4.0
;   SEQ ID NO 80
;   LENGTH: 401
;   TYPE: PRT
;   ORGANISM: OspC Chimera
US-09-596-746-80

Query Match   84.3%;   Score 1524.5; DB 19; Length 401;
Best Local Similarity   86.0%;   Pred. No. 4e-100;
Matches 326; Conservative 10; Mismatches 32; Indels 11; Gaps 4;

Qy   2 ACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEL-AKAI 60
           :|||||||||||||||||||||||||||||||||||||||||||||||||||: |||
Db   25 SCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAI 84

Qy   61 GKKIKNDGSLDNEANRNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEEFST 120
           ||| : ||| | |||||||||||||||| |||:| ||||| | |||||||||||||||: :
Db   85 GKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLK-NEGLKEKIDAAKKCSETFTN 143

Qy   121 KLKDNHAQLGIQGVTDENAKKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANSV 180
           ||| : || | :|||| |||||||||| || |||||||||||||||||||||||||||||
Db   144 KLKEKHTDLGKEGVTDADAKEAILKTNGT-KTRGAEELGKLFESVEVLSKAAKEMLANSV 202

Qy   181 KELTSPVHG--------NNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKE 232
           ||||||| |||||||||||||||||||||||||||||||||||||
Db   203 KELTSPVVAESPKKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKE 262

Qy   233 VETLLTSIDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELK 292
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   263 VETLLTSIDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELK 322

Qy   293 AEIEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVK 352
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   323 AEIEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVK 382

Qy   353 NLSKAAKEMLTNSVKELTS 371
           |||||||||||||||||||
Db   383 NLSKAAKEMLTNSVKELTS 401

Search completed: March 18, 2002, 10:08:45
Job time: 973 sec

OM protein - protein search, using sw model

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

Run on: March 18, 2002, 09:58:21 ; Search time 81.02 seconds
                (without alignments)
                292.523 Million cell updates/sec

Title: US-09-596-746A-34
Perfect score: 1809
Sequence: 1 MACNNSGKDGNTSANSADES.........KNLSKAAKEMLTNSVKELTS 371

Scoring table: BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched: 303921 seqs, 63882009 residues

Database :   Pending_Patents_AA_New:*
                1: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pep:*
                2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep:*
                3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep:*
                4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep:*
                5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep:*
                6: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep1:*
                7: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep:*
                8: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep:*

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Total number of hits satisfying chosen parameters: 303921

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match Length	DB	ID	Description
1	886	49.0	211	6	US-09-974-992-5	Sequence 5, Appli
2	656	36.3	207	6	US-09-974-992-3	Sequence 3, Appli
3	556	30.7	212	6	US-09-974-992-7	Sequence 7, Appli
4	187.5	10.4	1304	6	US-09-708-427-15045	Sequence 15045, A
5	187.5	10.4	1313	6	US-09-708-427-15044	Sequence 15044, A
6	187	10.3	837	6	US-09-815-242-5883	Sequence 5883, Ap
7	187	10.3	837	7	US-10-072-851-5883	Sequence 5883, Ap
8	187	10.3	875	7	US-10-072-851-13080	Sequence 13080, A
9	187	10.3	875	6	US-09-815-242-13080	Sequence 13080, A
10	182.5	10.1	1881	7	US-10-032-585-7646	Sequence 7646, Ap
11	182.5	10.1	1881	7	US-10-072-851-15590	Sequence 15590, A
12	176.5	9.8	1014	7	US-09-708-427-19883	Sequence 19883, A
13	176.5	9.8	1018	6	US-09-708-427-19882	Sequence 19882, A
14	176.5	9.8	1269	6	US-09-708-427-19881	Sequence 19881, A
15	169.5	9.4	1144	6	US-09-708-427-15046	Sequence 15046, A
16	168.5	9.3	730	6	US-09-708-427-3731	Sequence 3731, Ap
17	168.5	9.3	805	6	US-09-708-427-3730	Sequence 3730, Ap
18	168.5	9.3	841	6	US-09-708-427-3729	Sequence 3729, Ap
19	162.5	9.0	1690	6	US-09-614-150-10224	Sequence 10224, A
20	162.5	9.0	1690	6	US-09-614-150-10311	Sequence 10311, A
21	162.5	9.0	1833	6	US-09-614-150-00215	Sequence 00215, A
22	160	8.8	2434	6	US-10-072-851-5835	Sequence 5835, Ap
23	160	8.8	2434	7	US-10-072-851-5835	Sequence 5835, Ap
24	160	8.8	6281	6	US-09-815-242-12996	Sequence 12996, A
25	160	8.8	6281	7	US-10-072-851-12996	Sequence 12996, A
26	158	8.7	2346	6	US-09-614-150-17349	Sequence 17349, A
27	155.5	8.6	1948	7	US-10-032-585-7611	Sequence 7611, Ap
28	155.5	8.6	1948	7	US-10-072-851-15555	Sequence 15555, A
29	154	8.5	2437	7	US-09-815-242-5834	Sequence 5834, Ap
30	154	8.5	2437	7	US-10-072-851-5834	Sequence 5834, Ap
31	153.5	8.5	2056	6	US-09-614-150-4824	Sequence 4824, Ap
32	152.5	8.4	1786	6	US-09-742-006-3	Sequence 3, Appli
33	152	8.4	679	6	US-09-614-150-25386	Sequence 25386, A
34	151.5	8.4	3070	6	US-09-961-403-7	Sequence 7, Appli
35	151.5	8.4	3110	6	US-09-561-709B-7	Sequence 7, Appli
36	149.5	8.3	2344	7	US-10-072-851-12713	Sequence 12713, A
37	149.5	8.3	2344	7	US-09-815-242-12713	Sequence 12713, A
38	149	8.2	1639	6	US-09-614-150-6213	Sequence 6213, Ap
39	149	8.2	1409	6	US-09-815-242-23910	Sequence 23910, A
40	148	8.2	5795	6	US-10-072-851-12610	Sequence 12610, A
41	148	8.2	5795	7	US-09-815-242-12610	Sequence 12610, A
42	147	8.1	722	6	US-09-815-242-12888	Sequence 12888, A
43	147	8.1	722	7	US-10-072-851-12888	Sequence 12888, A
44	147	8.1	991	6	US-09-815-242-5803	Sequence 5803, Ap
45	147	8.1	991	7	US-10-072-851-5803	Sequence 5803, Ap

ALIGNMENTS

RESULT 1
US-09-974-992-5
Sequence 5, Application US/09974992
GENERAL INFORMATION:
; APPLICANT: Mathiesen, Marianne J.
; APPLICANT: Theisen, Michael
; APPLICANT: Holm, Arne
; APPLICANT: Ostergaard, Soren
; TITLE OF INVENTION: Novel OspC-derived peptide fragments
; FILE REFERENCE: 459-666P
; CURRENT APPLICATION NUMBER: US/09/974,992
; CURRENT FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: 09/180,089
; PRIOR FILING DATE: 1999-05-13
; PRIOR APPLICATION NUMBER: PCT/DK97/00203
; PRIOR FILING DATE: 1997-05-02
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 5
; LENGTH: 211
; TYPE: PRT
; ORGANISM: Borrelia burgdorferi
US-09-974-992-5

Query Match   49.0%; Score 886; DB 6; Length 211;
Best Local Similarity 97.3%; Pred. No. 5.9e-54;
Matches 183; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

```
Qy    1 MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDELAKAI 60
          :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   17 ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDELAKAI 76

Qy   61 GKKIKNDGSLDNEARNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAARKCSEEFST 120
         |||||||||| :||| ||||||||||||||||||||||||||||||||||||||||||
Db   77 GKKIKNDGSLGDEANNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAARKCSEEFST 136

Qy  121 KLKDNHAQLGIGQGVTDENAKKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANSV 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  137 KLKDNHAQLGIGQGVTDENAKKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANSV 196

Qy  181 KELTSPVV 188
         ||||||||
Db  197 KELTSPVV 204
```

RESULT 2
US-09-974-992-3

Sequence 3, Application US/09974992
GENERAL INFORMATION:
APPLICANT: Mathiesen, Marianne J.
APPLICANT: Theisen, Michael
APPLICANT: Holm, Arne
APPLICANT: Ostergaard, Soren
TITLE OF INVENTION: Novel OspC-derived peptide fragments
FILE REFERENCE: 459-666P
CURRENT APPLICATION NUMBER: US/09,974,992
CURRENT FILING DATE: 2001-10-10
PRIOR APPLICATION NUMBER: 09/180,089
PRIOR FILING DATE: 1999-05-13
PRIOR APPLICATION NUMBER: PCT/DK97/00203
PRIOR FILING DATE: 1997-05-02
NUMBER OF SEQ ID NOS: 40
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3
LENGTH: 207
TYPE: PRT
ORGANISM: Borrelia garinii

Query Match   36.3%;   Score 656;   DB 6;   Length 207;
Best Local Similarity   75.0%;   Pred. No. 3.8e-38;
US-09-974-992-3
Matches 141; Conservative 20; Mismatches 23; Indels 4; Gaps 3;

```
Qy   1 MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEVALLSIDELAKAI 60
         :|:|||| || |||:|||||||||||||||||||:|||:||||||||||||:|||||:||
Db  17 ISCNNSG--GDSASTNPDESAKGPNLTVISKKITDSNAFLLAVKEVEVALLSSIDELSKAI 74

Qy  61 GKKIKNDGSLDNEANRNESLLAGAYTISTLITQKLSKLINGSEGLKEKIAAAKKCSEEFST 120
         ||||||||:|||||||||||:|||| :|| ||||||:|    |||||||: |||||||:|
Db  75 GKKIKNDGTLDNEANRNESLIAGAYEISKLITQKLSVLN-SEELKEKIKEAKDCSEKFTI 133

Qy 121 KLKDNHAQLGIGQVTDENAKKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANSV 180
         ||||.|| ||| .|.||||||||||:| || ||||:|||| :||||||||||:| |||||||
Db 134 KLKDSHAFLGIQSVQDDNAKAILKTHGT-KDKGAKELEELFKSLESLSKAAQAALTNSV 192

Qy 181 KELTSPVV 188
         ||||:|||
Db 193 KELTNPVV 200
```

RESULT 3
US-09-974-992-7
Sequence 7, Application US/09974992
GENERAL INFORMATION:
APPLICANT: Mathiesen, Marianne J.
APPLICANT: Theisen, Michael
APPLICANT: Holm, Arne
APPLICANT: Ostergaard, Soren
TITLE OF INVENTION: Novel OspC-derived peptide fragments
FILE REFERENCE: 459-666P
CURRENT APPLICATION NUMBER: US/09,974,992
CURRENT FILING DATE: 2001-10-10
PRIOR APPLICATION NUMBER: 09/180,089
PRIOR FILING DATE: 1999-05-13
PRIOR APPLICATION NUMBER: PCT/DK97/00203
PRIOR FILING DATE: 1997-05-02
NUMBER OF SEQ ID NOS: 40
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 7
LENGTH: 212
TYPE: PRT
ORGANISM: Borrelia afzelii

Query Match   30.7%;   Score 555;   DB 6;   Length 212;
Best Local Similarity   67.2%;   Pred. No. 2.9e-31;
US-09-974-992-7
Matches 123; Conservative 19; Mismatches 39; Indels 2; Gaps 2;

RESULT 4
US-09-708-427-15045
Sequence 15045, Application US/09708427
GENERAL INFORMATION:
APPLICANT: N. ALEXANDROV et al.
TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
TITLE OF INVENTION: THEREBY
FILE REFERENCE: 2750-1243P
CURRENT APPLICATION NUMBER: US/09/708,427
CURRENT FILING DATE: 2000-11-09
NUMBER OF SEQ ID NOS: 85364
SOFTWARE: PatentIn version 3.1
SEQ ID NO 15045
LENGTH: 1304
TYPE: PRT
ORGANISM: Arabidopsis thaliana
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1..1304
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: 1..1304
OTHER INFORMATION: Ceres Seq. ID 1828628

Query Match   10.4%;   Score 187.5;   DB 6;   Length 1304;
Best Local Similarity   23.4%;   Pred. No. 7.3e-05;
US-09-708-427-15045
Matches 102; Conservative 75; Mismatches 149; Indels 109; Gaps 17;

```
Qy 191 NNSGKDGNT-SANSADESVKGPNLTEISKKITESNAVVLAVKEVEVETLLTSIDELA-KAIG 248
         |||| ||| :|:|| |||||||||||||||:||||:|||||||| |:|| :||||| ||||
Db  20 NNSGKGGDSASTNPDESAKGPNLTEISKKITDSNAFLLAVKEVEVLVLSIDELAKKAIG 79

Qy 249 KKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAEIEKAKKCSEEFTAK 308
         :|| |: ||||| |||||| |||| || ||| |:| | ||||:|||||||||||||| ||
Db  80 QKIDNNNGLAALNNQNGSLLAGAYAISTLITEKLSKLKNLEELKTEIAKAKKCSEEFINK 139

Qy 309 LKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLSKAAKEMLTNSVKE 368
         || || ||||  |||:|||:|||||  | ||| ||| ||||| ||||:||||||||||||
Db 140 LKSGHADLGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGLLKAAQVALTNSVKE 199

Qy 369 LTS 371
         |||
Db 200 LTS 202
```

Query Match   10.4%;   Score 187.5;   DB 6;   Length 1304;
Best Local Similarity   23.4%;   Pred. No. 7.3e-05;
Matches 102; Conservative 75; Mismatches 149; Indels 109; Gaps 17;

```
Qy   4 NNSGKDGNTSANSADES---------------VKGPNLTEISKKITDSNAVLLAVK------EVEA 48
         :|||| || :| ||                  ||||||||||||||:|||||||||       |:
Db  85 NEKLKEALAAQKRAEESFEVEKFRAVELEQAGLEAVQKRKDVTSKNELESIRSQHALDISA 144

Qy  49 LLSSIDELAKAIGKKIKNDGSLDNEA-NRNESLLAGAYTISTLITQK------------LSKLNG 100
         | ::||| ::|:|| |:||:|:|| | :|:|  ||||| |:||: |            |:|||
Db 145 LSTTEEL----QRVKHELSMTADAKNKALSHAEEATKIAEIHAEKAEIIASELGRLKA 199

Qy 101 SEGLKEKIAAAKKCSEEFSTKLKDNHAQLGIGQVTDENAKKAILKANAAGKDKGVEELE- 159
         : | ||| :|| |    |:  ||:||||| |||||||||||:||||:|  |:|||:|||:
Db 200 LLGSKEEKEAIE--GNEIVSKLKSE----IELLRGELEKVSILESSLKEQEGLVEQLKV 252

Qy 160 --KLSGSLESLSKAAKEMLANSVKELTSPVVHGNNSGKDGNTSANSADESVKGPNLTEIS 217
         |:||:||||:|:||||:| :|: |:|||   |||| ||| :||||| |||||||||:
Db 253 DLEAAKMAESCTNSSVEEWKNKVHELEKEV-----EESNRSKSSASESME----SVM 300

Qy 218 KKITESNAVV---------------LAVKEVETLLITSIDE-------LAKAIGKKIKNDV- 255
         :||| |||:               ||||||| |||::|||       ||||||||||||:
Db 301 KQLAELNHVLHETKSDNAAQREKIELLEKTIEAQRTDLEEYGRQVCIAKEEASKIENLVE 360

Qy 256 -------------------SLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAEIEKAKK 300
                            ||||| |||||| |||| || |||||:|:| | |||:|||||||
Db 361 SIKSELEISQEEKTRALDNEKAATSN-------IQNLLDQRTELSIELERCKVEEEKSKK 413

Qy 301 CSEEFTAKLKGEHTDLGKEGVTDNAKKAILKTNNDKTKGADELEKLFESVKNLSKAAKE 360
         |
```

```
Db    414 DMESLTLAL------QEASTESSEAKATLLVCQEELKNC---ESQVDSLKLASKETNE 462
Qy    361 ----MLTNSVKELTS 371
           ||:|  ||:|
Db    463 KYEKMLEDARNEIDS 477

RESULT 5
US-09-708-427-15044
; Sequence 15044, Application US/09708427
; GENERAL INFORMATION:
;  APPLICANT: N. ALEXANDROV et al.
;  TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
;  TITLE OF INVENTION: THEREBY
;  FILE REFERENCE: 2750-1243P
;  CURRENT APPLICATION NUMBER: US/09/708,427
;  CURRENT FILING DATE: 2000-11-09
;  NUMBER OF SEQ ID NOS: 85364
;  SOFTWARE: Patentin version 3.1
;
;  SEQ ID NO 15044
;  LENGTH: 1313
;  TYPE: PRT
;  ORGANISM: Arabidopsis thaliana
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: 1..1313
;  OTHER INFORMATION: Xaa is any amino acid
;  NAME/KEY: misc_feature
;  LOCATION: 1..1313
;  OTHER INFORMATION: Ceres Seq. ID 1828627

Query Match             10.4%;  Score 187.5;  DB 6;  Length 1313;
Best Local Similarity   23.4%;  Pred. No. 7.3e-05;
Matches 102;  Conservative 75;  Mismatches 149;  Indels 109;  Gaps 17;

Qy      4 NNSGKDGNTSANSADES----------VKGPNLTEISKKITDSNAVLLAVK------EVEA 48
            | ||                      |:||                     ::|
Db     94 NEKLKEALAAQKRAEESFEVEKFRAVELEQAGLEAVQKKDVTSKNELESIRSQHALDISA 153

Qy     49 LLSSIDELAKAIGKKIKNDGSLDNEA-NRNESLLAGAYTISTLLITQK-------LSKLNG 100
           |||:|| |:||:  |: |  :||:|  |: ||                     |:|
Db    154 LLSTTEEL----QRVKHELSMTADAKNKNALSHAEEATKIAEIHAEKAEILASELGRLKA 208

Qy    101 SEGLKEKIAAAKKCSEEFSTKLKDNHAQLGIQGVTDENAKKAILKANAAGKDKGVEELE- 159
            | |:|| |  |||  ||: |:|: :|     |  |||  || || |  | | || |
Db    209 LLGSKEEKEAIE--GNEIVSKLKSE----IELLRGELEKVSILESSLKEQEGLVEQLKV 261

Qy    160 --KLSGSLESLSKAAKEMLANSVKELTSPVVHGNNSGKDGNTSANSADESVKGPNLTEIS 217
             :||| |||: || ||:|| ||:|||||||||||||||| |||||||||:|||||:
Db    262 DLEAAKAAESCTNSSVEEWKNKVHELEKEV---EESNRSKSASESME-----SVM 309

Qy    218 KKITESNAVV-----LAVKEVETLLTSIDE------LARAIGKKIKNDV- 255
           ||||:| ||      | |:||   |::|:       | : |  |:||
Db    310 KQLAELNVLHETKSDNAQKEKIELLEKTIEAQRTDLEEYGRQVCIAKEEASKLENLVE 369

Qy    256 -------------SLDNEADHINGSLISGAYLISNLITKKISAIKDSGELKAEIEKAKK 300
                        |||:||  ||||| :|   ||  |: || ||:
Db    370 SIKSELEISQEEKTRALDNEKAATSN-------IQNLLDQRTELSIELERCKVEEEKSKK 422

Qy    301 CSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLSKAAKE 360
           |||||||||||||||||  |:||||||||||| |::| |||||:|:||||||:||||:
Db    423 DMESLTLAL-------QEASTESSEAKATLLVCQEELKNC---ESQVDSLKLASKETNE 471

Qy    361 ----MLTNSVKELTS 371
            ||:|  ||:|
Db    472 KYEKMLEDARNEIDS 486

RESULT 6
US-09-815-242-5883
; Sequence 5883, Application US/09815242
```

```
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT: Ohlsen, Karl L.
;  APPLICANT: Zyskind, Judith W.
;  APPLICANT: Wall, Daniel
;  APPLICANT: Trawick, John D.
;  APPLICANT: Carr, Grant J.
;  APPLICANT: Yamamoto, Robert J.
;  APPLICANT: Xu, H. Howard
;  APPLICANT: Yamamoto, Robert T.
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION: Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE: 2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
;
;  SEQ ID NO 5883
;  LENGTH: 837
;  TYPE: PRT
;  ORGANISM: Staphylococcus aureus
US-09-815-242-5883

Query Match             10.3%;  Score 187;  DB 6;  Length 837;
Best Local Similarity   22.2%;  Pred. No. 4.0e-05;
Matches 101;  Conservative 62;  Mismatches 183;  Indels 108;  Gaps 17;

Qy     11 NTSANSADESVKGPNLTEISKKITDS------NAVLLAVKEVEALLSSIDELAKAIGKKIK 65
           |:|                            ::|
Db    278 NTAMTNLKNGIQDQNTIKQGVNFTDADEAKRNAYTNAVTQAEQILN-----KAQGPNTS 331

Qy     66 NDG-----------SLDNEANRESLLAGAYTISTLLTQKLSKLNGS--EGLKEKIAAAKKC 114
           ||:           ||||: :|
Db    332 KDGVETFALENVQRAKNELNGNQN-VANAKTTAKNALNNLTSINNAQKEALKSQIEGATTV 390

Qy    115 S--EEFSTKLKD-NHAQLGIQ-GVTDENAKKAILKANAAGKDKGVEELEKLS-------G 163
               ||  |:|   |: |:|  ||||| || || |
Db    391 AGVNQVSTTASELNTAMSNLQNGINDEAATKAAL-------NGTQNLEKAKQHANTAID 442

Qy    164 SLESLSKAAKEMLANSVKELTSPVVHGNNSGKDGNTSA-------NSADESV-KGPNLT 214
           ||||                                     :
Db    443 GLSHLTNAQKEALKQLVQQSTTVAEAQGNEQKANNVDAAMDKLRQSIADNATTKQNQNYT 502

Qy    215 EIS--KKITESNAVV----------------------------------LAVKE 232
            |:  ||||:| ||                                   |:||
Db    503 DASQNKDAYNNAVTTAQGIIDQTTSPFLDPTVINQAAGQVSTTKNALNGNENLEAAKQQ 562

Qy    233 VETLLTSIDELAKAIGKKIKNDVSLDNEADHINGSLISGAYLISNLITKKISAIKDSGELK 292
           |  |::|::                       ::|
Db    563 ASQSLGSLDNLNNAQKQTVTDQINGAHTVDEANQIKQNAQNLNTAMGNLKQAIADKDATK 622

Qy    293 AEI-----EKAKK------CSEEFTAKLKGEHTDLG--KEGVTDDNAKKAILKTNNDK 337
           ::|     ||:|       |:
Db    623 ATVNFTDADQAKQQAYNTAVTNAENIISKANGGNATQAEVEQAIKQVNAAKQALNGNANV 682

Qy    338 TKGADELEKLFESVKNLSKAAKEMLTNSVKELTS 371
           ||||:|:|:|||||||:|
Db    683 QHAKDEATALINSSNDLNQAQKDALKQQVQNATT 716
```

RESULT 7
US-10-072-851-5883
Sequence 5883, Application US/10072851

; GENERAL INFORMATION:
; APPLICANT: Carr, Grant J.
; APPLICANT: Xu, H. Howard
; APPLICANT: Foulkes, J. Gordon
; APPLICANT: Zamudio, Carlos
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Roemer, Terry
; APPLICANT: Jiang, Bo
; APPLICANT: Boone, Charles
; APPLICANT: Bussey, Howard
; TITLE OF INVENTION: Methods for Identifying the Target of a Compound which Inhibits c
; FILE REFERENCE: ELITRA.028A
; CURRENT APPLICATION NUMBER: US/10,072,851
; CURRENT FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 15811
; SEQ ID NO 5883
; LENGTH: 837
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-072-851-5883

Query Match   10.3%; Score 187; DB 7; Length 837;
Best Local Similarity 22.2%; Pred. No. 4.4e-05;
Matches 101; Conservative 62; Mismatches 183; Indels 108; Gaps 17;

Qy   11 NTSANSADESVKGPNLTEISKKITDS-----NAVLLAVKEVEALLSSIDELAKAIGKKIK 65
           |:|:| :||: || || ::| || |||:|| |:: |: || ||:|| |
Db   278 NTAMTNLKNGIQDQNTIKQGVNFTDADEAKRNAYTNAVTQAEQILN-----KAQGPNTS 331

Qy   66 NDG--------SLDNEANRNESLLAGAYTISTLITQKLSKLNGS--EGLKEKIAAAKKC 114
           | | |: : ||:: | | | ::|:: | |: || || |:|
Db   332 KDGVETALENVQRAKNELNGNQN-VANAKTTAKNALNNLTSINNAQKEALKSQIEGATTV 390

Qy   115 S--EEFSTKLKD-NHAQLGIQ-GVTDENAKKAILKANAAGKDKGVEELEKLS-------G 163
           : |:| :| : | || || || :| :| || || |: |
Db   391 AGVNQVSTTASELNTAMSNLQNGINDEAATKAAL------NGTQNLEKAKQHANTAID 442

Qy   164 SLESLSKAAKEMLANSVKELTSPVVHGNNSGKDGNTSA-------NSADESV--KGPNLT 214
           |: | |:| |: ||:| :|| :| || : ||:|: ||||:
Db   443 GLSHLTNAQKEALKQLVQQSTTVAEAQGNEQKANNVDAAMDKLRQSIADNATTKQNQNYT 502

Qy   215 EIS--KKITESNAVV--LAVKE 232
           ::| |:|:| : |: :
Db   503 DASQNKKDAYNNAVTTAQGIIDQTTSPTLDPTVINQAAGQVSTTKNALNGNENLEAAKQQ 562

Qy   233 VETLLITSIDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELK 292
           |:|: |:|:: ||| |:| | | | | :| :| ::| | ||:||::|| |
Db   563 ASQSIGSLDNLNNAQKQTVTDQINGAHTVDEANQIKQNAQNLNTAMGNLKQAIADKDATK 622

Qy   293 AEI------EKAKK--------CSEEFTAKLKGEHTDLG--KEGVTDDNAKKAILKTNNDK 337
           |:| |||| :::||:|:||:|| ||||::|:|| || |||
Db   623 ATVNFTDADQAKQQAYNTAVTNAENIISKANGGNATQAEVEQAIKQVNAAKQALNGNANV 682

Qy   338 TKGADELEKLFESVKNLSKAAKEMLTNSVKELTS 371
           | |:|::|::|:|: |:|:|:| :| |||||
Db   683 OHAKDEATALINSSNDLNQAQKQDALKQQVQNATT 716

RESULT 8

US-09-815-242-13080
Sequence 13080, Application US/09815242

; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SEQ ID NO 13080
; LENGTH: 875
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-13080

Query Match   10.3%; Score 187; DB 6; Length 875;
Best Local Similarity 22.2%; Pred. No. 4.6e-05;
Matches 101; Conservative 62; Mismatches 183; Indels 108; Gaps 17;

Qy   11 NTSANSADESVKGPNLTEISKKITDS-----NAVLLAVKEVEALLSSIDELAKAIGKKIK 65
           |:|:| :||: || || ::| || |||:|| |:: |: || ||:|| |
Db   315 NTAMTNLKNGIQDQNTIKQGVNFTDADEAKRNAYTNAVTQAEQILN-----KAQGPNTS 368

Qy   66 NDG--------SLDNEANRNESLLAGAYTISTLITQKLSKLNGS--EGLKEKIAAAKKC 114
           | | |: : ||:: | | | ::|:: | |: || || |:|
Db   369 KDGVETALENVQRAKNELNGNQN-VANAKTTAKNALNNLTSINNAQKEALKSQIEGATTV 427

Qy   115 S--EEFSTKLKD-NHAQLGIQ-GVTDENAKKAILKANAAGKDKGVEELEKLS-------G 163
           : |:| :| : | || || || :| :| || || |: |
Db   428 AGVNQVSTTASELNTAMSNLQNGINDEAATKAAL------NGTQNLEKAKQHANTAID 479

Qy   164 SLESLSKAAKEMLANSVKELTSPVVHGNNSGKDGNTSA-------NSADESV--KGPNLT 214
           |: | |:| |: ||:| :|| :| || : ||:|: ||||:
Db   480 GLSHLTNAQKEALKQLVQQSTTVAEAQGNEQKANNVDAAMDKLRQSIADNATTKQNQNYT 539

Qy   215 EIS--KKITESNAVV--LAVKE 232
           ::| |:|:| : |: :
Db   540 DASQNKKDAYNNAVTTAQGIIDQTTSPTLDPTVINQAAGQVSTTKNALNGNENLEAAKQQ 599

Qy   233 VETLLITSIDELAKAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELK 292
           |:|: |:|:: ||| |:| | | | | :| :| ::| | ||:||::|| |
Db   600 ASQSIGSLDNLNNAQKQTVTDQINGAHTVDEANQIKQNAQNLNTAMGNLKQAIADKDATK 659

Qy   293 AEI------EKAKK--------CSEEFTAKLKGEHTDLG--KEGVTDDNAKKAILKTNNDK 337
           |:| |||| :::||:|:||:|| ||||::|:|| || |||
Db   660 ATVNFTDADQAKQQAYNTAVTNAENIISKANGGNATQAEVEQAIKQVNAAKQALNGNANV 719

Qy   338 TKGADELEKLFESVKNLSKAAKEMLTNSVKELTS 371
           | |:|::|::|:|: |:|:|:| :| |||||

Db   720 QHAKDEATALINSSNDLNQAQKDALKQQVQNATT 753

RESULT   9
US-10-072-851-13080
; Sequence 13080, Application US/10072851
; GENERAL INFORMATION:
; APPLICANT: Carr, Grant J.
; APPLICANT: Xu, H. Howard
; APPLICANT: Foulkes, J. Gordon
; APPLICANT: Zamudio, Carlos
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Roemer, Terry
; APPLICANT: Jiang, Bo
; APPLICANT: Boone, Charles
; APPLICANT: Bussey, Howard
; TITLE OF INVENTION: Methods for Identifying the Target of a Compound which Inhibits C
; TITLE OF INVENTION: Proliferation
; FILE REFERENCE: ELITRA.028A
; CURRENT APPLICATION NUMBER: US/10/072,851
; CURRENT FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 15811
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13080
; LENGTH: 875
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-072-851-13080

Query Match   10.3%; Score 187; DB 7; Length 875;
Best Local Similarity   22.2%; Pred. No. 4.6e-05;
Matches 101; Conservative 62; Mismatches 183; Indels 108; Gaps 17;

Qy   11 NTSANSADESVKGPNLTEISKKIDS----------NAVLLAVKEVEALLSSIDELAKAIGKKIK 65
Db   315 NTAWTNLKNGIQDQNTIKQCVNFTDADEARKNAYTNAVTQAEQILN--------KAQGPNTS 368

Qy   66 NDG---------SLDNEANRNESLLAGAYTISTLITQKLSKLNGS------EGLKEKIAAAKKC 114
Db   369 KDCGVETALENVQRAKNELNGNQN-VANAKTTAKNALNNLTSINNAQKEALKSQIEGATTV 427

Qy   115 S--EEFSTKLKD-NHAQLGIQ-GVTDENARKKAILKANAAGKDKGVEELEKLS----G 163
Db   428 AGVNQVSTTASELNTAMSNLQNGINDEAATKAAL------NGTQNLEKAKQHANTAID 479

Qy   164 SLESLSKAAKEMLANSVKELTSPVVHGNNSGKDGNTSA------NSADESV-KGPNLT 214
Db   480 GLSHLTNAQKEALKQLVQQSTTVAEAQGNEQKANNVDAAMDKLRQSIADNATTKQNQNYT 539

Qy   215 EIS--KKITESNAVV---------------LAVKE 232
Db   540 DASQNKKDAYNNAVTTAQGIIDQTTPSPTLDPTVINQAAGQVSTTKNALGNENLEAAAKQQ 599

Qy   233 VETLLTLTSIDELAKAIGKKIRNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELK 292
Db   600 ASQSLGSLDNLNNAQKQTVTDQINGAHTVDEANQIKQNAQNLNTAMGNLKQAIADKDATK 659

Qy   293 AEI------EKAKK-----CSEEFTAKLKGEHTDLG--KEGVTDDNAKKAILKTNNDK 337
Db   660 ATVNFTDADQAQAYNTAVTNAENIISKANGGNATQAEVEQAIKQVNAAKQALNGNANV 719

Qy   338 TKGADELEKLFESVKNLSKAAKEMLTNSVKELTS 371
Db   720 QHAKDEATALINSSNDLNQAQKDALKQQVQNATT 753

RESULT   10
US-10-032-585-7646
; Sequence 7646, Application US/10032585
; GENERAL INFORMATION:
; APPLICANT: Terry, Roemer D.
; APPLICANT: Bo, Jiang
; APPLICANT: Charles, Boone
; APPLICANT: Howard, Bussey
; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
; FILE REFERENCE: 10102-005-999
; CURRENT APPLICATION NUMBER: US/10/032,585
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 8000
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 7646
; LENGTH: 1881
; TYPE: PRT
; ORGANISM: Candida albicans
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (1881)..(1881)
; OTHER INFORMATION: X-any amino acid
US-10-032-585-7646

Query Match   10.1%; Score 182.5; DB 7; Length 1881;
Best Local Similarity   22.6%; Pred. No. 0.00026;
Matches 101; Conservative 75; Mismatches 139; Indels 131; Gaps 20;

Qy   26 LIEISKKIT---DSNAVLL-AVKEVEALL----------SSIDELAKAIGKKIKND 67
Db 1443 LTEATSELITKLQDNNQSLITEEIEKTKAALTKSSKDLEVCGNQKSELQDSLKSVKSELKN- 1501

Qy   68 GSLDNEANRNESLLAG-------AYTISTLITQKLSKL-------NGSEGLKEKIAAA 111
Db 1502 --FENKINQETTISLKDEIEEKQKEIVTLQQTELKDRISEVEKERAMLSENSETVIKEYSDK 1559

Qy   112 KKCSEEFSTKLKDNHAQLGIQGVTDENARKKAILKANAAGKDKGVEELEKLSGSLESLSKA 171
Db 1560 IKSLESKINSIKENHS----KEITTHNEQKTSLK--------ODIAKLSQDHES-AQT 1604

Qy   172 AKEMLANSVKELTSPVVHGNNSGKDGNTSANSADESVKGPNLTEISKKI----------TES 223
Db 1605 QLEDKENQLKELKASL-----EKHNTESATSIEE--KNNQIKELSETIKSLKTELKTSG 1656

Qy   224 NAVVLAVKEVETLLTLTSIDELAKAIGK------KIKNDVSLDNEADHNGSLISGAYLISNL 277
Db 1657 DALKQSOKEYKTLKTKNSDTESKLEKQLEELEKVKSDLQTADE-KLKGITEREIALKSEL 1715

Qy   278 ITKKISAIKDSGELKA------EIEKAK-------KCSEEFTAKLKGEHTDLGK-- 318
Db 1716 ETVKNSGLSTTSELAALTKTVKSLEKEKEELQFLSGNKSKELEDYIQK----HSDISEKL 1771

Qy   319 EGVTD--------DNAKKAILKTNNDKTKGADELE-------- 345
Db 1772 KALTDELEKTKQFDDSKKKLTELENDLTSTKKELETEKTQTSKFKNLEERKDKEIVKLN 1831

Qy   346 KLFESVKNLSKAAKEMLTNSVKELTS 371
Db 1832 KELELLKNDNSGAKKELSEKVSKLES 1857

RESULT   11
US-10-072-851-15590
; Sequence 15590, Application US/10072851
; GENERAL INFORMATION:
; APPLICANT: Carr, Grant J.
; APPLICANT: Xu, H. Howard
; APPLICANT: Foulkes, J. Gordon
; APPLICANT: Zamudio, Carlos
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.

; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Roemer, Terry
; APPLICANT: Jiang, Bo
; APPLICANT: Boone, Charles
; APPLICANT: Bussey, Howard
; TITLE OF INVENTION: Methods for Identifying the Target of a Compound which Inhibits ...
; FILE REFERENCE: ELITRA.028A
; CURRENT APPLICATION NUMBER: US/10/072,851
; CURRENT FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 15811
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15590
; LENGTH: 1881
; TYPE: PRT
; ORGANISM: Candida albicans
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (1881)..(1881)
; OTHER INFORMATION: X=any amino acid
; US-10-072-851-15590

Query Match   10.1%;   Score 182.5;   DB 7;   Length 1881;
Best Local Similarity   22.6%;   Pred. No. 0.00026;
Matches 101;   Conservative 75;   Mismatches 139;   Indels 131;   Gaps 20;

```
Qy    26  LTEISKKIT-------DSNAVLL-AVKEVEALL------------SSIDELAKAIGKKIKND  67
          ||| :||           : :: ||  | ||||                 ||  |  :|||
Db  1443  LTEATSELFKLQDDNNQSLIEEIEKTKAALTKSSKDLEVCGNQKSELQDSLKSVKSELKN- 1501

Qy    68  GSLDNEANRNESLLAG--------AYTISTLITQKLSKL--------NGSEGLKEKIAAA  111
          ||:                     : :| |  |                |:: | ||:
Db  1502  --FENKYNQETTSLKDEIEEKQKEIVTLQTELKDRISEVEKERAMLSENSETVIKEYSDK 1559

Qy   112  KKCSEEFSTKLKDNHAQLGIQGVTDENAKKAILKANAAGKDKGVEELEKLSGSLESLSKA  171
          :|  |  :|                                               |:
Db  1560  IKSLESKINSIKENHS----KEITTHNEQKTSLK------------QDIAKLSQDHES-AQT 1604

Qy   172  AKEMLANSVKELTSPVVHGNNSGKDGNTSANSADESVKGPNLTEISKKI--------TES  223
          ||: |||:|                                                  |:|
Db  1605  QLEDKENQLKELKASL-----EKHNTESATSIEE--KNNQIKELSETIKSLKTELKTSG 1656

Qy   224  NAVVLAVKEVETLLTSIDELAKAIGK------EIEKAK------KCSEEFTAKILKGEHTDLGK--  277
          :|  |                                      | ::|  |  |||
Db  1657  DALKQSOKEYKTLKNSDTESKLEQLEEELEVKVKSDLQTADE-KLKGITERELALKSEL 1715

Qy   278  ITKKISAIKDSGELKA------EIEKAK--------KCSEEFTAKILKGEHTDLGK--  318
          : |:                                  |::| | | |  |||
Db  1716  ETVKNSGLSTTSELAALTKTVKSLEKEKEELQFLSGNKSKELEDYIQK----HSDISEKL 1771

Qy   319  EGVTD--------DNAKKAILKTNNDKTKGADELE----------  345
                      |||::| | || ||
Db  1772  KALTDELKEKTKQFDDSKKLTELENDLTSTKKELETEKTQTSKFKNLEERKDKEIVKLN 1831

Qy   346  KLFESVKNLSKAAKEMLTNSVKELTS 371
          | : ::| || | |: | || |:
Db  1832  KELELKNDNSGAKKELSEKVSKLES 1857
```

RESULT 12
US-09-708-427-19883
; Sequence 19883, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1243P

; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19883
; LENGTH: 1014
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..1014
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..1014
; OTHER INFORMATION: Ceres Seq. ID 1836847
; US-09-708-427-19883

Query Match   9.88%;   Score 176.5;   DB 6;   Length 1014;
Best Local Similarity   24.4%;   Pred. No. 0.0003;
Matches 93;   Conservative 81;   Mismatches 152;   Indels 55;   Gaps 16;

```
Qy    16  SADESV--KGPNLTEISKKITDSNAV--LLAVKEVEALLSSIDEL------AKAIGKKIK  65
          | |:     |  | ||  :| :|         :  ||  | :: ||        ||| || |
Db   516  SAEESLEQKGREIDEATTKRMELEALHQSLSIDSEHRLQRAMEEFTSRDSEASSLTEKLR 575

Qy    66  NDGSLDNEANRNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEEFSTKLKDN 125
          :|  |      :|:   :  :| :       | ||:| |||   |||            |||
Db   576  D---LEGKIKSYEEQLAEASGCKSSSLKEKLEQTLG-----RLAAAESVNE---KLKQE 622

Qy   126  HAQLGIQGVTDENAKKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANSVKELTS 185
          | | ::     |  ::   : |  |:  | ||| |   :| | :| : | ||  :|
Db   623  FDQAQEKSLQSSSSESELLAETNNQLKIK-IQELEGLIGSGSVEKETALKRLEEAIERFNQ 681

Qy   186  PVVHGNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVL--AVKEVETLLTSIDEL 243
          |            :  | |::    | | :: :|: :|:  ||    ||  ::|    |||
Db   682  KETESSDLVERLKTHENQIEEYKK---LAHEASGVADTRKVELEDALSKIKNLESTIEEL 738

Qy   244  A------KAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKA-EI 295
          |        | |:    |:|  :|  :|:|||  : |:||   ||   ::  | |: |
Db   739  GAKCQGLEKESGDLAEVNLKINLELANHGS------EANELQTKLSALEAEKEQTANEL 791

Qy   296  EKAKKCSEEFTAKL--KGEHTDLGREGVTDDNAKKAILKTNNDKTKGADELEKLFESVK- 352
          | | |   |      |:   ||   :|| |:::|:  : |    :| || |:|| |:|
Db   792  EASKTTIEDLTKQLTSEGEKLQSQIEKLRAVAAEKSVLESH----FEELEKTLSEVKA 845

Qy   353  NLSKAAKEMLTNSVK--ELTS 371
          ||| |    ||  ||   ||||
Db   846  QLKENVENAATASVKAELTS 866
```

RESULT 13
US-09-708-427-19882
; Sequence 19882, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID...
; FILE REFERENCE: 2750-1243P

; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19882
; LENGTH: 1018
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..1018
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..1018

```
; OTHER INFORMATION: Ceres Seq. ID 1836846
US-09-708-427-19982

Query Match          9.8%;   Score 176.5;  DB 6;  Length 1018;
Best Local Similarity   24.4%;  Pred. No. 0.0003;
Matches   93;  Conservative  81;  Mismatches 152;  Indels  55;  Gaps  16;

Qy   16  SADESV--KGPNLTEISKKITDSNAV--LLAVKEVEALLSSIDEL-------AKAIGKKIK   65
         |:||:|  ||||||||:|||:|||    |:||                   ::||
Db  520  SAEESLEQKGREIDEATTKRMELEALHQSLSIDSEHRLQKAMEEFTSRDSEASSLTEKLR  579

Qy   66  NDGSLDNEANRNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEEFSTKLKDN  125
Db  580  D---LEGKIKSYEEQLAEASGKSSSLKEKLEQTLG------RLAAAESVNE----KLKQE  626

Qy  126  HAQLGIQGVTDENAKKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANSVKELTS  185
Db  627  FDQAQEKSLQSSSESELLAETNNQLKIK-IQELEGLIGSGSVEKETALKRLEBAIERFNQ  685

Qy  186  PVVHGNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVL--AVKEVETLLTSIDEL  243
Db  686  KETESSDLVEKLKTHENQIEEYKK---LAHEASGVADTRKVELEDALSKLNLESTIEEL   742

Qy  244  A------KAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKA-EI   295
Db  743  GAKCQGLEKESGDLAEVNLKNLELANHGS------EANELQTKLSALEAEKEQTANEL   795

Qy  296  EKAKKCSEEFTAKL--KGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVK-  352
Db  796  EASKTTIEDLTKQLTSEGEKLQSQIEKLRAVAAEKSVLESH------FEELEKTLSEVKA  849

Qy  353  NLSKAAKEMLTNSVK---ELTS  371
Db  850  QLKENVENAATASVKAELTS   870
```

```
RESULT 14
US-09-708-427-19981
; Sequence 19881, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: THEREBY
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: Patentin version 3.1
;
; SEQ ID NO 19881
; LENGTH: 1269
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..1269
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..1269
; OTHER INFORMATION: Ceres Seq. ID 1836845

Query Match          9.8%;   Score 176.5;  DB 6;  Length 1269;
Best Local Similarity   24.4%;  Pred. No. 0.0004;
Matches   93;  Conservative  81;  Mismatches 152;  Indels  55;  Gaps  16;

Qy   16  SADESV--KGPNLTEISKKITDSNAV--LLAVKEVEALLSSIDEL------AKAIGKKIK   65
Db  771  SAEESLEQKGREIDEATTKRMELEALHQSLSIDSEHRLQKAMEEFTSRDSEASSLTEKLR  830

Qy   66  NDGSLDNEANRNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEEFSTKLKDN  125
```

```
RESULT 15
US-09-708-427-15046
; Sequence 15046, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
; TITLE OF INVENTION: THEREBY
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: Patentin version 3.1
;
; SEQ ID NO 15046
; LENGTH: 1144
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..1144
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..1144
; OTHER INFORMATION: Ceres Seq. ID 1828629

Query Match          9.4%;   Score 169.5;  DB 6;  Length 1144;
Best Local Similarity   29.3%;  Pred. No. 0.0011;
Matches  110;  Conservative  42;  Mismatches 154;  Indels  69;  Gaps  18;

Qy   15  NSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDELAKAIGKKIKNDGSLDNEA   74
Db  606  NLVDNVANMQNIAEESKDLRE------REV-AYLKKIDELSTANGTLADNVTNLQNIS   656

Qy   75  NRNE------SLLAGAYTISTL---ITQKLSKL---NGSEGLKEKIAAAKKCSEEFST   120
Db  657  EENKELRERETTLLKKAEELSELNESLVDKASKLQTVVQENEELRERETAYLKKIEELSK  716

Qy  121  K---LKDNHAQLGIQGVTDENAKKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLA  177
Db  717  LHEILSDQETKLQI-----SNHEKEELKERETAYLKIEELSKV--------QEDLL    760

Qy  178  NSVKELTSPVVHGNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVV-----LAVKE  232
Db  761  NKENELHGMVEIEDL--RSKDSLAQKKIEELSNFNASLLIKE--NELQAVVCENEELKSKQ  818

Qy  233  VETLLTSIDELAKAIGKKIKNDVSLDNEADHNGSLISGAYL-----ISNLITKKISAIKDS  288
Db  819  VSTLKT-IDELSDLKQSLIHEKEKELQAAIVENEKLKAEAALSLQRIEELTNLKQTLIDKQ  877

Db  831  D---LEGKIKSYEEQLAEASGKSSSLKEKLEQTLG------RLAAAESVNE----KLKQE  877
Qy  126  HAQLGIQGVTDENAKKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANSVKELTS  185
Db  878  FDQAQEKSLQSSSESELLAETNNQLKIK-IQELEGLIGSGSVEKETALKRLEBAIERFNQ  936
Qy  186  PVVHGNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVL--AVKEVETLLTSIDEL  243
Db  937  KETESSDLVEKLKTHENQIEEYKK---LAHEASGVADTRKVELEDALSKLNLESTIEEL   993
Qy  244  A------KAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKA-EI   295
Db  994  GAKCQGLEKESGDLAEVNLKNLELANHGS------EANELQTKLSALEAEKEQTANEL   1046
Qy  296  EKAKKCSEEFTAKL--KGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVK-  352
Db 1047  EASKTTIEDLTKQLTSEGEKLQSQIEKLRAVAAEKSVLESH------FEELEKTLSEVKA 1100
Qy  353  NLSKAAKEMLTNSVK---ELTS  371
Db 1101  QLKENVENAATASVKAELTS  1121
```

```
Qy   289  GELKA------EIEKAKKCS--EEFTAKLKGEHTDLGKEG----VTDDNAKKAILKTNND-  336
           |||         ||||| | | |  ::  | :|| |: |
Db   878  NELQGVFHENEELKAKEASSLKKIDELLHLEQSWLEKESEFQRVTQENLE---LKTQDAL  934

Qy   337  KTKGADELEKLFESV  351
           | :|| ||  ||:
Db   935  AAKKIEELSKLKESL  949
```

OM protein - protein search, using sw model

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

Run on: March 18, 2002, 09:56:47 ; Search time 68.77 Seconds
                 (without alignments)
                 410.946 Million cell updates/sec

Title: US-09-596-746A-34
Perfect score: 1809
Sequence: 1 MACNNSGKDGNTSANSADES..........KNLSKAAKEMLTNSVKELTS 371

Scoring table: BLOSUM62
                 GapOP 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Maximum DB seq length: 0
Minimum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   PIR_68:*
                 1: pir1:*
                 2: pir2:*
                 3: pir3:*
                 4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	884	48.9	193	2	S70280	outer surface prot
2	883	48.8	211	2	S69918	outer surface prot
3	879	48.8	210	2	I40144	outer surface prot
4	862	47.7	203	2	I40108	outer surface prot
5	849	46.9	178	2	I40106	outer surface prot
6	663.5	36.7	210	2	S69927	outer surface prot
7	662	36.6	193	2	S70287	outer surface prot
8	658.5	36.4	210	2	G70218	outer surface prot
9	656	36.3	200	2	I40122	outer surface prot
10	656	36.3	207	2	S69919	outer surface prot
11	654	36.2	193	2	S70286	outer surface prot
12	653	36.1	207	2	S37727	outer surface prot
13	651.5	36.0	193	2	S70276	outer surface prot
14	649.5	35.9	212	2	I40143	outer surface prot
15	647	35.8	207	2	I40271	outer surface prot
16	643	35.5	189	2	S70263	outer surface prot
17	640	35.4	191	2	I40153	outer surface prot
18	638	35.3	200	2	S54198	outer surface prot
19	636.5	35.2	211	2	I40268	outer surface prot
20	636	35.2	209	2	S69917	outer surface prot
21	635	35.1	207	2	I40276	outer surface prot
22	628	34.7	191	2	S70284	outer surface prot
23	622	34.4	211	2	I40145	outer surface prot
24	621	34.3	193	2	S70279	outer surface prot
25	615.5	34.0	194	2	S70277	outer surface prot
26	613	33.9	192	2	S70285	outer surface prot
27	613	33.9	193	2	S70259	outer surface prot
28	612	33.8	193	2	S69922	outer surface prot
29	610	33.7	176	2	I40121	outer surface prot
30	609.5	33.7	190	2	S70273	outer surface prot
31	608	33.6	210	2	S69923	outer surface prot
32	608	33.6	212	2	I40279	outer surface prot
33	607.5	33.6	177	2	I40129	outer surface prot
34	607	33.6	194	2	S70268	outer surface prot
35	606	33.5	178	2	I40125	outer surface prot
36	605.5	33.5	193	2	S70274	outer surface prot
37	603	33.3	178	2	I40104	outer surface prot
38	602.5	33.3	199	2	S54197	outer surface prot
39	599	33.1	209	2	I40273	outer surface prot
40	597	33.0	191	2	S70278	outer surface prot
41	595.5	32.9	211	2	S69932	outer surface prot
42	592	32.7	209	2	I40281	outer surface prot
43	590.5	32.6	214	2	S69916	outer surface prot
44	589.5	32.6	194	2	S70289	outer surface prot
45	588	32.5	209	2	I40142	outer surface prot

ALIGNMENTS

RESULT   1
S70280
outer surface protein C - Lyme disease spirochete
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 26-May-2000
C;Accession: S70280
R;Livey, I.; Gibbs, C.P.; Schuster, R.; Dorner, F.
Mol. Microbiol. 18, 257-269, 1995
A;Title: Evidence for lateral transfer and recombination in OspC variation in Lyme di
A;Reference number: S70255; MUID:96296448
A;Accession: S70280
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-193 <LIV>
A;Cross-references: EMBL:L42868; NID:g858735; PIDN:AAB37011.1; PID:g1695226
A;Experimental source: strain ZS7
C;Genetics:
A;Gene: ospC
C;Superfamily: Lyme disease spirochete surface protein C

Query Match   48.9%; Score 884; DB 2; Length 193;
Best Local Similarity   98.4%; Pred. No. 2.9e-34;
Matches 183; Conservative 1; Mismatches   2; Indels   0; Gaps   0;

Qy   3 CNNSGKDGNTSANSADESVKGPNLTEISKKTTDSNAVLLAVKEVEALLSSIDELAKAIGK 62
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 CNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDELAKAIGK 60

Qy   63 KIKNDGSLDNEANRNESLLAGAYTISTLITGKLSKLNGSEGLKEKIAAAKKCSEEFSTKL 122
            ||||||| ||| :|||||||||||||||||||||||||||||||||||||||||||||||
Db   61 KIKNDGSLGDEANHNESLLAGAYTISTLITGKLSKLNGSEGLKEKIAAAKKCSEEFSTKL 120

Qy   123 KDNHAQLGIQGVTDENAKKAIIKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANSVKE 182
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 KDNHAQLGIQGVTDENAKKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANSVKE 180

Qy   183 LTSPVV 188
            ||||||
Db   181 LTSPVV 186

RESULT   2
S69918
outer surface protein C precursor - Lyme disease spirochete (strain PBre)
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Variety: strain PBre
C;Date: 06-Dec-1996 #sequence_revision 14-Feb-1997 #text_change 26-May-2000
C;Accession: S69918; S72674; I40103
R;Jauris-Heipke, S.; Liegl, G.; Preac-Mursic, V.; Roessler, D.; Schwab, E.; Soutschek
J. Clin. Microbiol. 33, 1860-1866, 1995
A;Title: Molecular analysis of genes encoding outer surface protein C (OspC) of Borre

A;Reference number: I40047; MUID:95395018
A;Accession: S69918
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-211 <JAD>
A;Cross-references: EMBL:X81522; NID:g872021; PIDN:CAA57242.1; PID:g872022
A;Experimental source: strain PBre
R;Roessler, D.
submitted to the EMBL Data Library, September 1994
A;Reference number: S72674
A;Accession: S72674
A;Molecule type: DNA
A;Residues: 1-152,'E',154-211 <ROE>
A;Cross-references: EMBL:X81522; NID:g872021; PIDN:CAA57242.1; PID:g872022
C;Genetics:
A;Gene: ospc
C;Superfamily: Lyme disease spirochete surface protein C

Query Match   48.8%; Score 883; DB 2; Length 211;
Best Local Similarity 96.8%; Pred. No. 3.5e-34;
Matches 182; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

```
Qy   1 MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEVALLSSIDELAKAI  60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  17 ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEVALLSSIDELAKAI  76

Qy  61 GKKIKNDGSLDNEANRNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEEFST 120
         ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  77 GKKIKNDGSLGDEANHNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEEFST 136

Qy 121 KLKDNHAQLGIQGVTDENAKKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANSV 180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 137 KLKDNHAQLGIQGVTDDNAKKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANSV 196

Qy 181 KELTSPVV 188
         ||||||||
Db 197 KELTSPVV 204
```

RESULT 3
I40144
outer surface protein C precursor - Lyme disease spirochete
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 21-Jul-2000
C;Accession: I40144; S70282
R;Theisen, B.; Barthold, S.W.
Stevenson, B.; J. Bacteriol.
A;Title: Expression and sequence of outer surface protein C among North American isolate...
A;Reference number: I40143; MUID:95154673
A;Accession: I40144
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-210 <RES>
A;Cross-references: EMBL:U04281; NID:g434663; PIDN:AAC43297.1; PID:g434664
R;Livey, I.; Gibbs, C.P.; Schuster, R.; Dorner, F.
Mol. Microbiol. 18, 257-269, 1995
A;Title: Evidence for lateral transfer and recombination in OspC variation in Lyme disea...
A;Reference number: S70282
A;Accession: S70255; MUID:96296448
A;Molecule type: DNA
A;Residues: 19-210 <LIV>
A;Cross-references: EMBL:L42893; NID:g858721; PIDN:AAB37001.1; PID:g1695218
A;Experimental source: strain 297
C;Genetics:
A;Gene: ospC
C;Superfamily: Lyme disease spirochete surface protein C

Query Match   48.6%; Score 879; DB 2; Length 210;
Best Local Similarity 99.4%; Pred. No. 5.3e-34;
Matches 180; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy 191 NNSGKDGNISANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLTSIDELAKAIGKK 250
         ||||||||:||||||||||||||||||||||:|||||||||||||:||||||||||||||
Db  20 NNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEVETLLTSIDELAKAIGKK  79

Qy 251 IKNDVSLDNEADHNGSLISGAYLIISNLITKKISAIKDSGELKAEIEKAKKCSEEFTAKIK 310
         |||||||||||||||||||||| ||| |||||||||||||||||||||||||||||| ||
Db  80 IKNDVSLDNEADHNGSLISGAYLISTLITKKISAIKDSGELKAEIEKAKKCSEEFTAKLK 139

Qy 311 GEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLSKAAKEMLTNSVKELT 370
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 140 GEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLSKAAKEMLTNSVKELT 199

Qy 371 S 371
         |
Db 200 S 200
```

RESULT 4
I40108
outer surface protein C - Lyme disease spirochete (strain DK7) (fragment)
C;Species: Borrelia burgdorferi (Lyme disease spirochete) (fragment)
C;Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 26-May-2000
C;Accession: I40108; S34176
R;Theisen, M.; Frederiksen, B.; Lebech, A.M.; Vuust, J.; Hansen, K.
J. Clin. Microbiol. 31, 2570-2576, 1993
A;Title: Polymorphism in ospC gene of Borrelia burgdorferi and immunoreactivity of Os...
A;Reference number: I40105; MUID:94075528
A;Accession: I40108
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-203 <RES>
A;Cross-references: EMBL:X73625; NID:g313275; PIDN:CAA52004.1; PID:g313276.
C;Genetics:
A;Gene: ospC
C;Superfamily: Lyme disease spirochete surface protein C

Query Match   47.7%; Score 862; DB 2; Length 203;
Best Local Similarity 95.7%; Pred. No. 3.1e-33;
Matches 179; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

```
Qy   1 MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEVALLSSIDELAKAI  60
         :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  17 ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEVALLSSIDELAKAI  76

Qy  61 GKKIKNDGSLDNEANRNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEEFST 120
         ||||||||||| ||| :|||||||||||||||||||||||||||||||||||||||||
Db  77 GKKIKNDGSLGDEANHNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEEFST 136

Qy 121 KLKDNHAQLGIQGVTDENAKKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANSV 180
         ||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db 137 KLKDNHAQLGIQGVTDENAKKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANSV 196

Qy 181 KELTSPV 187
         | | |
Db 197 KSLQSVV 203
```

RESULT 5
I40106
outer surface protein C - Lyme disease spirochete (strain HUR) (fragment)
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 26-May-2000
C;Accession: I40106; S54185
R;Theisen, M.; Borre, M.; Mathiesen, M.J.; Mikkelsen, B.; Lebech, A.M.; Hansen, K.
J. Bacteriol. 177, 3036-3044, 1995
A;Title: Evolution of the Borrelia burgdorferi outer surface protein OspC.
A;Reference number: I40104; MUID:95286481
A;Accession: I40106
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-178 <RES>

Query Match   47.8%; Score 862; DB 2; Length 203;
Best Local Similarity 95.7%; Pred. No. 3.1e-33;
Matches 179; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

```
Qy   1 MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEVALLSSIDELAKAI  60
         :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  17 ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEVALLSSIDELAKAI  76

Qy  61 GKKIKNDGSLDNEANRNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEEFST 120
         ||||||||||| ||| :|||||||||||||||||||||||||||||||||||||||||
Db  77 GKKIKNDGSLGDEANHNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEEFST 136

Qy 121 KLKDNHAQLGIQGVTDENAKKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANSV 180
         ||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db 137 KLKDNHAQLGIQGVTDENAKKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANSV 196

Qy 181 KELTSPV 187
         | | |
Db 197 KSLQSVV 203
```

C;Superfamily: Lyme disease spirochete surface protein C

A:Cross-references: EMBL:X84765; NID:g793789; PIDN:CAA59236.1; PID:g793790
C:Genetics:
A:Gene: ospC
C:Superfamily: Lyme disease spirochete surface protein C

```
Query Match  46.9%;  Score 849;  DB 2;  Length 178;
Best Local Similarity  99.4%;  Pred. No. 1e-32;
Matches 177;  Conservative 1;  Mismatches 0;  Indels 0;  Gaps 0;

Qy  11 NTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDELAKAIGKKIKNDGSL 70
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 NTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDELAKAIGKKIKNDGSL 60

Qy  71 DNEANRNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEEFSTKLKDNHAQLG 130
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  61 DNEANRNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEEFSTKLKDNHAQLG 120

Qy 131 IQGVTDENAKKAILKANNAAGKDKGVEELEKLSGGSLESLSKAAKEMLANSVKELTSPVV 188
        ||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db 121 IQGVTDENAKKAILEANAAGKDKGVEELEKLSGGSLESLSKAAKEMLANSVKELTSPVV 178
```

RESULT 6
S69927
outer surface protein C precursor - Lyme disease spirochete (strain Pka)
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
A:Variety: strain Pka
A:Date: 15-Feb-1997 #sequence_revision 27-Feb-1997 #text_change 26-May-2000
C:Accession: S69927; S72669
R:Jauris-Heipke, S.; Liegl, G.; Preac-Mursic, V.; Roessler, D.; Schwab, E.; Soutschek, E
J:Clin. Microbiol. 33, 1860-1866, 1995
A:Reference number: 140047; MUID:95395018
A:Title: Molecular analysis of genes encoding outer surface protein C (OspC) of Borrelia
C:Genetics:
C:Superfamily: Lyme disease spirochete surface protein C
A:Reference number: S72669
A:Accession: S72669
A:Molecule type: DNA
A:Residues: 1-210 <JAU>
A:Cross-references: EMBL:X69589
A:Experimental source: strain Pka
A:Residues: 1-124,'D',126-139,'E',141-210 <JAW>
submitted to the EMBL Data Library, February 1994
R:Jauris, S.

```
Query Match  36.7%;  Score 663.5;  DB 2;  Length 210;
Best Local Similarity  77.2%;  Pred. No. 3.8e-24;
Matches 146;  Conservative 10;  Mismatches 30;  Indels 3;  Gaps 3;

Qy   1 MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEL-AKA 59
         :|||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db  17 ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA 76

Qy  60 IGKKIKNDGSLDNEANRNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEEFS 119
        |||||:  | | : | : ||||||:|||||:|||| |:| ||||||| |:|||||::|:
Db  77 IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLK-NEGLKEKIEAAKKCSETFT 135

Qy 120 TKLKDNHAQLGIQGVTDENAKKAILKANNAAGKDKGVEELEKLSGGSLESLSKAAKEMLANS 179
         :|||: | :|| | |||:| |:||| |:   ||:| ||| ||  |:|:||||||||||||
Db 136 NKLKDKHFDLGKECVTDADAKEAILKTNGT-KTRGAEELGKLFESVEVLSKAAKEMLANS 194

Qy 180 VKELTSPVV 188
        |||||||||
Db 195 VKELTSPVV 203
```

RESULT 7
S70287
outer surface protein C - Lyme disease spirochete
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
A:Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 26-May-2000
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Reference number: S70255; MUID:96296448
R:Livey, I.; Gibbs, C.P.; Schuster, R.; Dorner, F.
Mol. Microbiol. 18, 257-269, 1995
A:Accession: S70287
A:Title: Evidence for lateral transfer and recombination in OspC variation in Lyme di
A:Residues: 1-193 <LIV>
A:Cross-references: EMBL:L42895; NID:g858723; PIDN:AAB37003.1; PID:g1695220
A:Experimental source: strain 28354
C:Genetics:
A:Gene: ospC
C:Superfamily: Lyme disease spirochete surface protein C

```
Query Match  36.6%;  Score 662;  DB 2;  Length 193;
Best Local Similarity  75.9%;  Pred. No. 4.1e-24;
Matches 142;  Conservative 17;  Mismatches 26;  Indels 2;  Gaps 2;

Qy   3 CNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDELA-KAIG 61
        |||||||||||||||||||||||||||||||:|||||||||||:|:||:|||||| |:||
Db   1 CNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVLLAVKEIETLLASIDELATRAIG 60

Qy  62 KKIKNDGSLDNEANRNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEEFSTK 121
        |||:: :|  : | : |:||||||||| |||||||  |: |:||||||::|||||:|:||
Db  61 KKIQONGGLAVYEAGHNGTLLAGAYTISKLITQKLDGLKNSERLKEKIENAKKCSEDFTK 120

Qy 122 LKDNHAQLGIQGVTDENAKKAILKANNAAGKDKGVEELEKLSGGSLESLSKAAKEMLANSVK 181
        |:  |||||| | |||:||||||| :: | |||| : ||||| :  |:|:||||||||||||
Db 121 LEGEHAQLGIENVTDENAKKAILIFDAA-KDKGAAELEKLFKAVENLAKAAKEMLANSVK 179

Qy 182 ELTSPVV 188
        |||||:|
Db 180 ELTSPIV 186
```

RESULT 8
G70218
outer surface protein C - Lyme disease spirochete
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C:Date: 13-Feb-1998 #sequence_revision 17-Feb-1998 #text_change 21-Jul-2000
A:Cross-references: GB:AE000792; NID:g3253098; PIDN:AAC66329.1; PID:g2689901; TIGR:BB
A:Experimental source: strain B31
A:Accession: G70218; I40269; S37726; S70281
R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; Wh
son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vu
; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997
A:Authors: Smith, H.O.; Venter, J.C.
A:Title: Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi.
A:Reference number: A70100; MUID:98065943
A:Accession: G70218
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-210 <KLE>
A:Cross-references: GB:AE000792; NID:g3253098; PIDN:AAC66329.1
A:Experimental source: strain B31
A:Reference number: 140269
J:Clin. Microbiol. 33, 2415-2420, 1995
R:Fukunaga, M.; Hamase, A.
A:Title: Outer surface protein C gene sequence analysis of Borrelia burgdorferi sensu
A:Reference number: 140269; MUID:96025162
A:Accession: I40269
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-210 <RES>
A:Cross-references: GB:D49497; NID:g7070092; PIDN:BAA08457.1; PID:g769684
R:Jauris-Heipke, S.; Fuchs, R.; Motz, M.; Preac-Mursic, V.; Schwab, E.; Soutschek, E.
Med. Microbiol. Immunol. 182, 37-50, 1993
A:Title: Genetic heterogenity of the genes coding for the outer surface protein C (Os

A;Reference number: S37726; MUID:93268136
A;Accession: S37726
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-210 <JAU>
A;Cross-references: EMBL:X69596; NID:g311391; PIDN:CAA49306.1; PID:g311392
R;Livey, I.; Gibbs, C.P.; Schuster, R.; Dorner, F.
Mol. Microbiol. 18, 257-269, 1995
A;Reference number: S70281
A;Accession: S70281
A;Reference number: S70255; MUID:96296448
A;Title: Evidence for lateral transfer and recombination in OspC variation in Lyme disea
C;Genetics:
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 19-210 <LIV>
A;Cross-references: EMBL:L42887; NID:g858715; PIDN:AAB36995.1; PID:g1695212
A;Experimental source: strain Ip2
C;Genetics:
Gene: ospC
Superfamily: Lyme disease spirochete surface protein C

```
Query Match           36.4%;     Score 658.5;  DB 2;   Length 210;
Best Local Similarity 76.7%;     Pred. No. 6.5e-24;
Matches  145;  Conservative 11;  Mismatches 30;  Indels 3;  Gaps 3;

Qy    1 MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEL-AKA 59
          |||:|||||||||||| ||||||||||||||||||||||||||||||||||||||| |||
Db   17 ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAKA 76

Qy   60 IGKKIKNDGSLDNEANRNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEEFS 119
          ||||||||| |||||||||||||||| ||||||:|||| | |||||||:|||||||:||
Db   77 IGKKIHONNGLDTENNINGSLLAGAYAISTLIKQKLDGLK-NEGLKEKIDAAKKCSETFT 135

Qy  120 TKIKDNHAQLGIQGVTDENAKKAILKANAAGKDKGVEELEKLSGGSLESLSKAAKEMLANS 179
          :| | ||| ||||| |||| |||| :||||| ||||||| || |:||:||||||||||||
Db  136 NKLKEKHTDLGKEGVTDADAKEAILKTNGT-KTKGAEELGKLFESVEVLSKAAKEMLANS 194

Qy  180 VKELTSPVV 188
          |||||:|||
Db  195 VKELTSPVV 203
```

RESULT 9
I40122
outer surface protein C - Lyme disease spirochete (strain DK6) (fragment)
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 26-May-2000
Accession: I40122; S34177
R;Theisen, M.; Frederiksen, B.; Lebech, A.M.; Vuust, J.; Hansen, K.
J. Clin. Microbiol. 31, 2570-2576, 1993
A;Title: Polymorphism in ospC gene of Borrelia burgdorferi and immunoreactivity of OspC
A;Reference number: I40105; MUID:9407528
A;Accession: I40122
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-200 <RES>
A;Cross-references: EMBL:X73626; NID:g313277; PIDN:CAA52005.1; PID:g313278
C;Genetics:
C;Gene: ospC
C;Superfamily: Lyme disease spirochete surface protein C

```
Query Match           36.38%;    Score 656;   DB 2;   Length 200;
Best Local Similarity 75.0%;     Pred. No. 8e-24;
Matches  141;  Conservative 20;  Mismatches 23;  Indels 4;  Gaps 3;

Qy    1 MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDELAKAI 60
          |::|||| ||| || | |||||||||:|||||||||||:|||||||||||||||||||||
Db   17 ISCNNSG--GDSASTNPDESAKGPNLTVISKKITDSNAFLLAVKEVEALLSSIDELAKAI 74

Qy   61 GKKIKNDGSLDNEANRNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEEFST 120
          ||||||||:|||||||||||:|||||:|:||||||| || | ||||||:|| |:|||:||
Db   75 GKKIKNDGTLDNEANRNESLIAGAYEISKLITQKLSVLN-SEELKEKIKEAKDCSEKFTT 133

Qy  121 KIKDNHAQLGIQGVTDENAKKAILKANAAGKDKGVEELEKLSGGSLESLSKAAKEMLANS 180
          ||||:||:||||.| || ||||||||:  | ||||| || | ||:|||||||| | :||
Db  134 KIKDSHAELGIQSVQDDNAKKAILKTHGT-KDKGAKELEELFKSLESLSKAAQAALTNSV 192

Qy  181 KELTSPVV 188
          ||||:|||
Db  193 KELTNPVV 200
```

RESULT 10
S69919
outer surface protein C precursor - Borrelia garinii (strain PTrob)
C;Species: Borrelia garinii
A;Variety: strain PTrob
A;Date: 06-Dec-1996 #sequence_revision 14-Feb-1997 #text_change 26-May-2000
A;Cross-references: EMBL:X83554; NID:g872027; PIDN:CAA58544.1; PID:g872028
A;Experimental source: strain PTrob
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1995
C;Genetics:
A;Gene: ospC
C;Superfamily: Lyme disease spirochete surface protein C
A;Accession: S69919
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-207 <JAU>
R;Jauris-Heipke, S.; Liegl, G.; Preac-Mursic, V.; Roessler, D.; Schwab, E.; Soutschek
J. Clin. Microbiol. 33, 1860-1866, 1995
A;Reference number: I40047; MUID:95395018
A;Title: Molecular analysis of genes encoding outer surface protein C (OspC) of Borre
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-207/Product: outer surface protein C #status predicted <MAT>

```
Query Match           36.38%;    Score 656;   DB 2;   Length 207;
Best Local Similarity 75.0%;     Pred. No. 8.3e-24;
Matches  141;  Conservative 20;  Mismatches 23;  Indels 4;  Gaps 3;

Qy    1 MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDELAKAI 60
          |::|||| ||| || | |||||||||:|||||||||||:|||||||||||||||||||||
Db   17 ISCNNSG--GDSASTNPDESAKGPNLTVISKKITDSNAFLLAVKEVEALLSSIDELAKAI 74

Qy   61 GKKIKNDGSLDNEANRNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEEFST 120
          ||||||||:|||||||||||:|||||:|:||||||| || | ||||||:|| |:|||:||
Db   75 GKKIKNDGTLDNEANRNESLIAGAYEISKLITQKLSVLN-SEELKEKIKEAKDCSEKFTT 133

Qy  121 KLKDNHAQLGIQGVTDENAKKAILKANAAGKDKGVEELEKLSGGSLESLSKAAKEMLANSV 180
          ||||:||:||||.| || ||||||||:  | ||||| || | ||:|||||||| | :||
Db  134 KLKDSHAELGIQSVQDDNAKKAILKTHGT-KDKGAKELEELFKSLESLSKAAQAALTNSV 192

Qy  181 KELTSPVV 188
          ||||:|||
Db  193 KELTNPVV 200
```

RESULT 11
S70286
outer surface protein C - Lyme disease spirochete
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 26-May-2000
A;Reference number: S70286
A;Accession: S70286
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-193 <LIV>
R;Livey, I.; Gibbs, C.P.; Schuster, R.; Dorner, F.
Mol. Microbiol. 18, 257-269, 1995
A;Reference number: S70255; MUID:96296448
A;Title: Evidence for lateral transfer and recombination in OspC variation in Lyme di
A;Cross-references: EMBL:L42897; NID:g858728; PIDN:AAB37006.1; PID:g1695222
A;Experimental source: strain 26815
C;Genetics:

A;Gene: ospC
C;Superfamily: Lyme disease spirochete surface protein C

Query Match   74.5%;   Score 654;   DB 2;   Length 193;
Best Local Similarity   74.5%;   Pred. No. 9.4e-24;
Matches 140;   Conservative 19;   Mismatches 25;   Indels 4;   Gaps 3;

```
Qy     3 CNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDELAKAIGK 62
           ||||||||| ||||||||||||||||| |||:|||||:|||||||:||||||||||||||
Db     1 CNNSGKDGNASTNSADESVKGPNLTEISKRITESNAVVLAVKEVAALLSSIDELAKAIGK 60

Qy    63 KIKNDGSLDNEANRNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKCSEEFSTKL 122
          || ||| ::|   | ||||||||:||  ||:|||  :|||||||||:  |:||::|| ||
Db    61 KINNNG-LDDVQNFNASLLAGAHTISKLVTEKLSLKKNSEGLKEKIEDARKKCSDDFTKKL 119

Qy   123 KDNHAQLGIGG--VTDENAKKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANSV 180
          :  ||||| ||  ||||:||||||:|||:  ||||:||||| ||:|||| ||| ||||||
Db   120 QSSHAQLGVAGGATTDEEAKKAILRTNAI-KDKGADELEKLFKSVESLAKAAQDALANSV 178

Qy   181 KELTSPVV 188
          :|||||||
Db   179 NELTSPVV 186
```

RESULT 12
S37727
outer surface protein C precursor - Lyme disease spirochete
A;Accession: S37727
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
A;Status: preliminary
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 26-May-2000
A;Molecule type: DNA
A;Residues: 1-207 <JAU>
A;Cross-references: EMBL:X69595; NID:g311393; PIDN:CAA49305.1; PID:g311394
C;Superfamily: Lyme disease spirochete surface protein C
R;Jauris-Helpke, S.; Fuchs, R.; Motz, M.; Preac-Mursic, V.; Schwab, E.; Soutschek, E.; W
R;Med. Microbiol. Immunol. 182, 37-50, 1993
A;Title: Genetic heterogenity of the genes coding for the outer surface protein C (ospC)
A;Reference number: S37727
A;Accession: S37726; MUID:9326136

Query Match   36.2%;   Score 653;   DB 2;   Length 207;
Best Local Similarity   36.1%;   Pred. No. 1.1e-23;
Matches 140;   Conservative 21;   Mismatches 23;   Indels 4;   Gaps 3;

```
Qy     1 MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDELAKAI 60
          : |||||  | :::::  |:|:|||||| ||||||||||  |||||||||||||||:|||
Db    17 ISCNNSG--GDSASTNPDESAKGPNLTVISKKITDSNAFLLAVKEVEALLSSIDELSKAI 74

Qy    61 GKKIKNDGSLDNEANRNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKCSEEFST 120
          |||||||| :|||||||||:|||||:||:||||||| :| ||:|||||: || ||:|:|
Db    75 GKKIKNDGTLDNEANRNESLIAGAYEISKLITQKLSVIN-SEELKEKIKEAKDCSQKFTT 133

Qy   121 KLKDNHAQLGIQGYTDENAKKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANSV 180
          |||| || ||||| :|:|||||||| |  ||||| :|| |:  ||||| ||| |||||||
Db   134 KLKDSHAELGIQSVQDDNAKKAILKTHGT-KDKGAKELELFKSLESLSRAAQAALTNSV 192

Qy   181 KELTSPVV 188
          ||||:|||
Db   193 KELTNPVV 200
```

RESULT 13
S70276
outer surface protein C - Lyme disease spirochete
A;Accession: S70276
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 26-May-2000
A;Accession: S70276
R;Livey, I.; Gibbs, C.P.; Schuster, R.; Dorner, F.
R;Mol. Microbiol. 18, 257-269, 1995
A;Title: Evidence for lateral transfer and recombination in OspC variation in Lyme disea

A;Reference number: S70255; MUID:96296448
A;Accession: S70276
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-193 <LIV>
A;Cross-references: EMBL:L42890; NID:g858718; PIDN:AAB36998.1; PID:g1695215
A;Experimental source: strain E61
C;Genetics:
A;Gene: ospC
C;Superfamily: Lyme disease spirochete surface protein C

Query Match   36.0%;   Score 651.5;   DB 2;   Length 193;
Best Local Similarity   76.5%;   Pred. No. 1.0e-23;
Matches 140;   Conservative 17;   Mismatches 23;   Indels 3;   Gaps 3;

```
Qy   191 NNSGKDG--NTSANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLTSIDELA-KAIG 248
          |||| ||  :::::  |:|:|||||| |||||||:|||||||||||||| |||||| |||
Db     2 NNSGKGGDSTSTNPADESAKGPNLTEISKKITDSNAFLLAVKEVEVTLVASIDELATKAIG 61

Qy   249 KKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAEIEKAKKCSEEFTAK 308
          ||||||  :||||:||||| :||| ||  ||:|:| :  |:||||:|  |||||::|| |
Db    62 KKIKNDGTLDNEANHNGSLLAGAYAISTLITQKLSVL-NSEELKAEIVKAKKCSEDFTKK 120

Qy   309 LKGEHTDLGKEGVTDNAKKAILKTNNDKTKGADELEKLFESVKNLSKAAKEMLTNSVKE 368
          || :||:||::  :|:||||||||| ||| || ||||| |||:.|||||||  ||||||
Db   121 LKDKHTELGKQDANDDAKKAILKTNGDKTLGAAELEKLSESVTSLSKAAKESLTNSVKE 180

Qy   369 LTS 371
          |||
Db   181 LTS 183
```

RESULT 14
I40143
outer surface protein C precursor - Lyme disease spirochete
A;Accession: I40143
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 21-Jul-2000
A;Accession: I40143
R;Stevenson, B.; Barthold, S.W.
R;FEMS Microbiol. Lett. 124, 367-372, 1994
A;Title: Expression and sequence of outer surface protein C among North American isol
A;Reference number: I40143; MUID:95154673
A;Accession: I40143
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-212 <RES>
A;Cross-references: EMBL:U04280; NID:g2314879; PIDN:AAC45542.1; PID:g434662
C;Genetics:
A;Gene: ospC
C;Superfamily: Lyme disease spirochete surface protein C

Query Match   35.9%;   Score 649.5;   DB 2;   Length 212;
Best Local Similarity   74.3%;   Pred. No. 1.7e-23;
Matches 142;   Conservative 19;   Mismatches 25;   Indels 5;   Gaps 4;

```
Qy     1 MACNNSGKDGNT-SANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDELAKA 59
          : ||||||||| :::::  |:|:|||||||||||||||||  ||||||| ||||||||||
Db    17 ISCNNSGKDGNSASTNPADESAKGPNLTEISKKITDSNAIVLAVKEVEVTLLSSIDELAKA 76

Qy    60 IGKKIKNDGSLDNEANRNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKCSEEFS 119
          ||||| ||| :|   |:||:||:||:||:| |:||||  :|||:|||||  ||||::|
Db    77 IGKKINNNG-LDVLQNFNASLLGGAHTISKLIITEKLSLKLNGSEELKEKIEAAKKCSDDFT 135

Qy   120 TKLKDNHAQLGIGG--VTDENAKKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLA 177
          .||| || ||| ||  ::|||||||||:|  :|||||:||| |:  ||||| |||||||
Db   136 KKLQSSHAELGVAGGATTDENAKKAILKSN-ADKTKGADELGKLFESVESLAKAAKEMLA 194

Qy   178 NSVKELTSPVV 188
          |||||||||||
Db   195 NSVKELTSPVV 205
```

```
RESULT  15
I40271
outer surface protein C precursor - Borrelia garinii
C;Species: Borrelia garinii
C;Date: 04-Sep-1997 #sequence_revision 04-Sep-1997 #text_change 26-May-2000
C;Accession: I40271
R;Fukunaga, M.; Hamase, A.
J. Clin. Microbiol. 33, 2415-2420, 1995
A;Title: Outer surface protein C gene sequence analysis of Borrelia burgdorferi sensu la
A;Reference number: I40269; MUID:96025162
A;Accession: I40271
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-207 <RES>
A;Cross-references: GB:D49377; NID:g1041103; PIDN:BAA08375.1; PID:g1041104
C;Superfamily: Lyme disease spirochete surface protein C

Query Match          35.8%;  Score 647;  DB 2;  Length 207;
Best Local Similarity   72.3%;  Pred. No. 2.1e-23;
Matches  136;  Conservative  27;  Mismatches  21;  Indels   4;  Gaps   3;

Qy       1 MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDELAKAI  60
             ::|||||  |:|:: : || ||||| |||||||||:||||||||:|:||||||||:||||
Db      17 ISCNNSG--GDTASTNPDESAKGPNLIEISKKITDSNAVVLAVKEVEALLSSIDELSKAI  74

Qy      61 GKKIKNDGSLDNEANRNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEEFST 120
            ||||||| ||||||||||||||||||| ||||||||| | | |||||| |||||:
Db      75 GKKIRNDGTLDNEANRNESLIAGAYEISKLITQKLSVLN-SEELKEKIKEAKDCSEKFTT 133

Qy     121 KLKDNHAQLGIQGVTDENAKKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANSV 180
            |||:|::||:||:|| ||:|||||||||||| : |||| ||| :|||:::|||: |||
Db     134 KLRDSHAELGVQNVQDDNAKRAILKTH-GNKDKGAKELKELSESLENLAKAAQAASSNSV 192

Qy     181 KELTSPVV 188
           |||||||
Db     193 KELTSPVV 200

Search completed: March 18, 2002, 09:56:48
Job time: 371 sec
```

OM protein - protein search, using sw model

Copyright (c) 1993 - 2000 Compugen Ltd.

GenCore version 4.5

Run on: March 18, 2002, 10:11:49 ; Search time 39.62 Seconds
                 (without alignments)
                 343.328 Million cell updates/sec

Title: US-09-596-746a-34
Perfect score: 1809
Sequence: 1 MACNNSGKDGNTSANSADES............KNLSKAAKEMLTNSVKELTS 371

Scoring table: BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   SwissProt_39.*

Post-processing:

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	658.5	36.4	210	1	OSC1_BORBU	Q07337; borrelia bu
2	556	30.7	212	1	OSC2_BORBU	O08137 borrelia bu
3	392.5	21.7	214	1	VM24_BORHE	P32778 borrelia he
4	373	20.6	215	1	VM03_BORHE	O02648 borrelia he
5	181	10.0	1957	1	YD86_SCHPO	Q10411 schizosach
6	180.5	10.0	1251	1	RBP2_PLAVB	Q00799 plasmodium
7	178.5	9.9	1790	1	USO1_YEAST	P25386 saccharomyc
8	173	9.6	564	1	M12_STRPY	P19401 streptococc
9	171.5	9.5	1509	1	MYSN_ACACA	P05659 acanthamoeb
10	169.5	9.4	775	1	YHGE_BACSU	P32399 bacillus su
11	169.5	9.4	1947	1	MYSC_CAEEL	P12845 caenorhabdi
12	168	9.3	1937	1	MYH8_HUMAN	P13535 homo sapien
13	167	9.2	1966	1	MYSB_CAEEL	P02566 caenorhabdi
14	166.5	9.2	1713	1	LM A3_HUMAN	O16787 homo sapien
15	165.5	9.1	2869	1	RBP1_PLAVB	Q00798 plasmodium
16	165	9.1	998	1	SCA4_RICAK	Q9aix9 rickettsia
17	164.5	9.1	296	1	OSB1_BORBU	P17739 borrelia bu
18	164.5	9.1	1938	1	MYH4_RABIT	Q28641 oryctolagus
19	163	9.0	1427	1	REST_HUMAN	P30622 homo sapien
20	163	9.0	1940	1	MYH3_CHICK	P02565 gallus gall
21	162.5	9.0	483	1	M6_STRPY	P08089 streptococc
22	162.5	9.0	1433	1	REST_CHICK	O42184 gallus gall
23	162.5	9.0	3210	1	CENF_HUMAN	P49454 homo sapien
24	161.5	8.9	1526	1	MYS2_SCHPO	Q9usi6 schizosacch
25	161.5	8.9	1941	1	MYH2_HUMAN	Q9ukx2 homo sapien
26	161	8.9	981	1	SCA4_RICFE	Q9aj37 rickettsia
27	161	8.9	1084	1	MYSS_RABIT	P02562 oryctolagus
28	160.5	8.9	1938	1	MYSS_CHICK	P13538 gallus gall
29	160	8.8	679	1	YKR9_YEAST	P34237 saccharomyc
30	159	8.8	1934	1	MYH7_MESAU	P13540 mesocricetu
31	159	8.7	1935	1	MYH7_HUMAN	P12883 homo sapien
32	157.5	8.7	1939	1	MYH1_HUMAN	P12882 homo sapien
33	157	8.7	1186	1	SMC_BACSU	P51834 bacillus su

34	157	8.7	1935	1	MYH7_PIG	P79293 sus scrofa
35	155.5	8.6	2349	1	TPR_HUMAN	P12270 homo sapien
36	155	8.6	1102	1	MYSC_CHICK	P29616 gallus gall
37	154.5	8.5	465	1	MYH6_RABIT	P04460 oryctolagus
38	154.5	8.5	2017	1	MYSN_DROME	Q99323 drosophila
39	154.5	8.5	2116	1	MYS2_DICDI	P08799 dictyosteli
40	153.5	8.5	1167	1	CAGA_HELPJ	Q9zit1 helicobacte
41	153	8.5	705	1	YNP9_CAEEL	P34562 caenorhabdi
42	153	8.5	1013	1	SCA4_RICRH	Q9aj81 rickettsia
43	153	8.5	1935	1	MYH7_RAT	P02564 rattus norv
44	151.5	8.4	539	1	MYS3_HYDAT	P39922 hydra atten
45	151.5	8.4	1939	1	MYH4_HUMAN	Q9y623 homo sapien

ALIGNMENTS

RESULT 1
OSC1_BORBU

ID	OSC1_BORBU	STANDARD;	PRT;	210 AA.

AC   Q07337;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   OUTER SURFACE PROTEIN C PRECURSOR (PC).
OS   OSPC OR BBB19.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OG   Plasmid lp54.
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 35210 / B31;
RX   MEDLINE=93268136; PubMed=8098841;
RA   Jauris-Heipke S., Fuchs R., Motz M., Preac-Mursic V., Schwab E.,
RA   Will G., Wilske B.;
RT   "Genetic heterogenity of the genes coding for the outer surface
RT   protein C (OspC) and the flagellin of Borrelia burgdorferi.";
RL   Med. Microbiol. Immunol. 182:37-50(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 35210 / B31;
RX   MEDLINE=93293332; PubMed=8478108;
RA   Wilske B., Preace-Mursic V., Jauris S., Pradel I., Soutschek E.,
RA   Schwab E., Wanner G.;
RT   "Immunological and molecular polymorphisms of OspC, an immunodominant
RT   major outer surface protein of Borrelia burgdorferi.";
RL   Infect. Immun. 61:2182-2191(1993).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 35210 / B31;
RX   MEDLINE=94041630; PubMed=8225587;
RA   Padula S.J., Sampieri A., Dias F., Szczepanski A., Ryan R.W.;
RT   "Molecular characterization and expression of p23 (OspC) from a North
RT   American strain of Borrelia burgdorferi.";
RL   Infect. Immun. 61:5097-5105(1993).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 35210 / B31;
RX   MEDLINE=96025162; PubMed=7494039;
RA   Fukunaga M., Hamase A.;
RT   "Outer surface protein C gene sequence analysis of Borrelia
RT   burgdorferi sensu lato isolates from Japan.";
RL   J. Clin. Microbiol. 33.2415-2420(1995).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 35210 / B31;
RX   MEDLINE=98055943; PubMed=9403685;
RA   Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA   Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
RA   Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
RA   Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,
RA   van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,

Column 1

```
RA   Utterback T., Watthey L., McDonald L., Artiach P., Bowman C.,
RA   Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
RA   Smith H.O., Venter J.C.;
RT   "Genomic sequence of a Lyme disease spirochaete, Borrelia
RT   burgdorferi.";
RL   Nature 390:580-586(1997).
CC   -!-  FUNCTION: NOT KNOWN; MAJOR IMMUNODOMINANT PROTEIN.
CC   -!-  SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID
CC        ANCHOR.
```

```
DR   EMBL; X69596; CAA49306.1; -.
DR   EMBL; U01894; AAA16058.1; -.
DR   EMBL; D49497; BAA08457.1; -.
DR   EMBL; AE000792; AAC66329.1; -.
DR   InterPro; IPR001800; -.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
DR   PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KW   Outer membrane; Lipoprotein; Signal; Plasmid; Antigen;
KW   Complete proteome.
FT   SIGNAL      1     18
FT   CHAIN      19    210       OUTER SURFACE PROTEIN C.
FT   LIPID      19     19       N-ACYL DIGLYCERIDE (BY SIMILARITY).
SQ   SEQUENCE   210 AA;  22340 MW;  7A4FC978F91777BF CRC64;
```

```
RESULT 2
ID   OSC2_BORBU     STANDARD;      PRT;   212 AA.
AC   Q08137;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   OUTER SURFACE PROTEIN C PRECURSOR (PC).
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
RX   MEDLINE=92219995; PubMed=1560779;
RA   Fuchs R., Jauris S., Lottspeich F., Preac-Mursic V., Wilske B.,
```

Column 2

```
RA   Soutschek E.;
RT   "Molecular analysis and expression of a Borrelia burgdorferi gene
RT   encoding a 22 kDa protein (pC) in Escherichia coli.";
RL   Mol. Microbiol. 6:503-509(1992).
RN   [2]
RP   SEQUENCE OF 1-205 FROM N.A.
RX   STRAIN=DK26;
RC   MEDLINE=94075528; PubMed=8253951;
RA   Theisen M., Frederiksen B., Lebech A.M., Vuust J., Hansen K.;
RT   "Polymorphism in ospC gene of Borrelia burgdorferi and
RT   immunoreactivity of OspC protein: implications for taxonomy and for
RT   use of OspC protein as a diagnostic antigen.";
RL   J. Clin. Microbiol. 31:2570-2576(1993).
CC   -!-  FUNCTION: NOT KNOWN; MAJOR IMMUNODOMINANT PROTEIN.
CC   -!-  SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID
CC        ANCHOR.
```

```
DR   EMBL; X62162; CAA44093.1; -.
DR   EMBL; X73624; CAA52003.1; -.
DR   InterPro; IPR001800; -.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
KW   Outer membrane; Lipoprotein; Signal; Plasmid; Antigen.
FT   SIGNAL      1     18
FT   CHAIN      19    212       OUTER SURFACE PROTEIN C.
FT   LIPID      19     19       N-ACYL DIGLYCERIDE (BY SIMILARITY).
SQ   SEQUENCE   212 AA;  22499 MW;  C206C231FBF2E7D4 CRC64;
```

```
Query Match       30.7%;  Score 556;  DB 1;  Length 212;
Best Local Similarity   67.2%;  Pred. No. 1.7e-18;
Matches 123;  Conservative 19;  Mismatches 39;  Indels 2;  Gaps 2;

Qy  191  NNSGKDGNT-SANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLTSIDELA-KAIG  248
         |||||:|| |:|||||||||||||||||||:||||:||||||||||||:|||||| |:|
Db   20  NNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETLVLSIDELAKKAIG   79

Qy  249  KKIKNDVSLDNEADHNGSLISGAYLISNIITIKKISAIKDSGELKAEIEKARKCSEEFTAK 308
         :|| |:  :||  |||||:|||||:|||||:|||||||:::|||||||||||||||||||
Db   80  QKIDNNNGLAALNNQNGSLLAGAYAISTLITEKLSKLKNEELKTEIAKKCSEEFTNK    139

Qy  309  LKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLSKAAKEMLTNSVKE  368
         || :|||||:  ||||:|||||||||:||:  |||||||||||:||||||||||||||||
Db  140  LKSGHADLGKQDATDDHAKAAILKTHATTDKGAXEFKDLFESVEGLLKRAAQAVALTNSVKE 199

Qy  369  LTS  371
         |||
Db  200  LTS  202
```

```
RESULT 3
ID   VM24_BORHE     STANDARD;      PRT;   214 AA.
AC   P32778;
DT   01-OCT-1993 (Rel. 27, Created)
DT   01-OCT-1993 (Rel. 27, Last sequence update)
DT   01-FEB-1994 (Rel. 28, Last annotation update)
DE   VARIABLE MAJOR OUTER MEMBRANE LIPOPROTEIN 24 PRECURSOR.
GN   VMP24.
OS   Borrelia hermsii.
OG   Plasmid.
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=140;
RN   [1]
RP   SEQUENCE FROM N.A.
```

Column 3

```
Query Match       36.4%;  Score 658.5;  DB 1;  Length 210;
Best Local Similarity   76.7%;  Pred. No. 5.7e-23;
Matches 145;  Conservative 11;  Mismatches 30;  Indels 3;  Gaps 3;

Qy    1  MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVERALLSSIDEL-AKA   59
         :||||||||||||||||||||||||||||||||||||||:|||||||||||||||| |||
Db   17  ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVILAVKEVERALLSSIDEIAAKA   76

Qy   60  IGKKIKNDGSLDNEARNESLLAGAYTISTLITQRLSKLNGSEGELKEKIAAAKKCSEEFS   119
         ||||| :|| ||::|:| |||||||:|||||:|:|  :|   |||||||:||||||:    135
Db   77  IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGIK-NEGLKEKIDAAKKCSETFI   135

Qy  120  TRLKRDNIAQLGLGIQGVTDENAKRAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANS  179
         |::|:::  :||:  |||||:||:|||||:| : |.|| |||.||:.|:|:|:|||||||||
Db  136  NKLKEKHTDLGKEGVTDADAKEAILKTNGT-KTKGAEELGKLFESSVEVLSKAAKEMLANS   194

Qy  180  VKELTSPVV  188
         |||||||||
Db  195  VKELTSPVV  203
```

RC STRAIN=SSP. HS1 SEROTYPE 24;
RX MEDLINE=93133110; PubMed=1484486;
RA Restrepo B.I., Kitten T., Carter C.J., Infante D., Barbour A.G.;
RT "Subtelomeric expression regions of Borrelia hermsii linear plasmids
RT are highly polymorphic.";
RL Mol. Microbiol. 6:3299-3311(1992).
CC -!- FUNCTION: SERVES TO AVOID THE HOST IMMUNE RESPONSE BY CHANGING
CC   FROM ONE SURFACE EXPOSED VMP TO ANOTHER.
CC -!- SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID
CC   ANCHOR.
CC -!- SIMILARITY: STRONG, TO VMP3.
CC --
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC --
DR EMBL; L04786; AAA22964.1; -.
DR InterPro; IPR001800; Lipoprotein_6.
Pfam; PF01441; Lipoprotein_6; 1.
ProDom; PD001149; Lipoprotein_6; 1.
PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KW Outer membrane; Lipoprotein; Signal; Plasmid.
FT SIGNAL   1   18
FT CHAIN   19   214   VARIABLE MAJOR OUTER MEMBRANE
FT   LIPOPROTEIN 24.
FT LIPID   19   19   N-ACYL DIGLYCERIDE (PROBABLE).
SQ SEQUENCE   214 AA; 22541 MW; F1583F510246F7C7 CRC64;

Query Match   21.7%;   Score 392.5; DB 1; Length 214;
Best Local Similarity   47.6%;   Pred. No. 2.4e-11;
Matches   90; Conservative 32; Mismatches 54; Indels 13; Gaps 3;

Qy 191 NNSGKDGNTSANSADESVKGP-----NLTEISKKITESNAVVLAVKEVETLLTSIDELAKA 246
           | | :::|| ||||| | ||:   :|| ||| |||| ||||:| |||| |||||||||
Db   20 NNGGPE------LKSDEVAKSDGTVLDLAKVSKKIKEASAFAASVKEVETLVKSVDELAKA   74

Qy 247 IGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKD----SGELKAEIEKAKKCS 302
           |||||||| | |:||||||||||:||||||||||||||| |||:|||:||| :
Db   75 IGKKIKNDGGLDTEAGQNGSLIAGVHSVSVSAVKIKVGALETTSGISNELKTITEVKSKA 134

Qy 303 EEFTAKLKGEHTDLGKEGVTDDNAKKRAILKTNNDKTKGADELEKLFESVKNLSKAAKEML 362
           | | | |||| :|| | ||| | | |:| || | ||| ||| |||:|| ||:|||
Db 135 EAFLNKLKDGHTELGKKDASDDDTKRAIKKDNSDKTKGASELEALNTAVDALLKAAEGEV 194

Qy 363 TNSVKELTS 371
           ::|||:|::
Db 195 EAAIKELTA 203

RESULT 4
ID VM03_BORHE   STANDARD;   PRT;   215 AA.
AC Q02448;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE VARIABLE MAJOR OUTER MEMBRANE LIPOPROTEIN 3 PRECURSOR.
GN VMP3.
OS Borrelia hermsii.
OG Plasmid.
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=140;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SSP. HS1 SEROTYPE 3;
RX MEDLINE=93133110; PubMed=1484486;
RA Restrepo B.I., Kitten T., Carter C.J., Infante D., Barbour A.G.;
RT "Subtelomeric expression regions of Borrelia hermsii linear plasmids

---

RT are highly polymorphic.";
RL Mol. Microbiol. 6:3299-3311(1992).
CC -!- FUNCTION: SERVES TO AVOID THE HOST IMMUNE RESPONSE BY CHANGING
CC   FROM ONE SURFACE EXPOSED VMP TO ANOTHER.
CC -!- SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID
CC   ANCHOR.
CC -!- SIMILARITY: STRONG, TO VMP24.
CC --
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC --
DR EMBL; L04789; AAA22967.1; -.
DR InterPro; IPR001800; Lipoprotein_6.
Pfam; PF01441; Lipoprotein_6; 1.
ProDom; PD001149; Lipoprotein_6; 1.
PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KW Outer membrane; Lipoprotein; Signal; Plasmid.
FT SIGNAL   1   18
FT CHAIN   19   215   VARIABLE MAJOR OUTER MEMBRANE
FT   LIPOPROTEIN 3.
FT LIPID   19   19   N-ACYL DIGLYCERIDE (PROBABLE).
SQ SEQUENCE   215 AA; 23139 MW; 684C74D35F87C771 CRC64;

Query Match   20.6%;   Score 373; DB 1; Length 215;
Best Local Similarity   46.3%;   Pred. No. 1.7e-10;
Matches   88; Conservative 32; Mismatches 56; Indels 14; Gaps 4;

Qy 191 NNSGKDGNTSANSADESVKGP-----NLTEISKKITESNAVVLAVKEVETLLTSIDELAKA 246
           | | :::|| ||||| | ||:   :|| ||| |||| ||||:| |||| |||||||||
Db   20 NNGGPE------LKSDEVAKSDGTVLDLAKISKKIKDASDFAASVKEVHTLVKSIDELAKA   74

Qy 247 IGKKIKNDVS-LDNEADHNGSLISGAYLISNLITKKISAIKD----SGELKAEIEKAKKC 301
           |||||||| | | |:|| ||| | ||:||||||||||||| |||:|||:|||
Db   75 IGKKIKNDNSNFEDENDHNGSLIAGVFQVILTVKAKLTSLEQIIGISDELKTEVGMVKKE 134

Qy 302 SEEFTAKLKGEHTDLGKEGVTDDNAKKRAILKTNNDKTKGADELEKLFESVKNLSKAAKEM 361
           || | | |||| :|| | ||| | | |:| || | ||| ||| |||:|| ||:|||
Db 135 SEAAYTQVKSKHTDLAKEGVTDAHRKSAILVTDGTKDKGAAAELIKLNTAIDELLKRANDA 194

Qy 362 LTNSVKELTS 371
           ::|||:|:::
Db 195 VETVIKELTA 204

RESULT 5
ID YD86_SCHPO   STANDARD;   PRT;   1957 AA.
AC Q10411;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE HYPOTHETICAL 222.8 KDA PROTEIN C1F3.06C IN CHROMOSOME I.
GN SPAC1F3.06C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomyces.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Connor R., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;
RL Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.
CC --
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

Left record

DR   Z706090; CAA94624.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   1957 AA; 222785 MW; 3F480CA0617LD9DA CRC64;

Query Match   10.0%; Score 181; DB 1; Length 1957;
Best Local Similarity   23.8%; Pred. No. 0.47;
Matches 110; Conservative   74; Mismatches 149; Indels 130; Gaps 23;

```
Qy    8 KDGNTSANSADE--------SVKGPNLTEISKKITDSNAVLLAV-KEVEALLSSIDE------- 55
          ||:  |    |  : ||:      ||  : ||  |: |||
Db  360 KDSRTNSQLEEMVELKESNRT-IHSQLTDAESKLSSFEQQENKSLKGSIDEYQNNLSSK 418

Qy   56 -------------LAKAIGK-----------KIKND---GSLDNEA 74
                      |:|:||           |:|:    ||:||:
Db  419 DKMVKQVSSQLEEARSSLAHATGKLAEINSERDFQNKKIKDFEKIEQDLRACLNSSSNEL 478

Qy   75 NRNESLLIAGAYTISTLLTQKLSKLNGSEGLKEKIAAAKKCSEEFSTKL----------KD 124
           ||:  ||  |  || |:||  |  | ::| |:     |   |        ::|  :KD
Db  479 KEK----------SALIDKKQELN--NLREQIKEQKKVSESTQSSLQSLQRDILNEKK 525

Qy  125 NH------AQLG-IQG-----VTDENAKKAILKANAAGKDKGV--EELEKLSGSLESLSKA 171
          :|       |:|  |:      :||:|:||||| :|  ||   |  |:|:|:||:|||:||:
Db  526 KHEVYEESQLNELKGELQTEINSEHLSSQLSTLAAEKEAAVATNNELSESKNSLQTLCNA 585

Qy  172 AKEMLANSVKELTSPVVHGNNSGKDGNTSANSADESVKGPNLTEISKKITESN-AVVLAV 230
           :|::| |:| : |  :|   :|  |: ||: | :  |  :| :|||:|| ||  |  ||
Db  586 FOEKLAKSVMQL------KENEQNFSSLDTSFK--KLNESHQELENNHQTITKQL 632

Qy  231 KEVETLLTSIDELAKAIGKKIKNDVSLDNE------ADHNGSLISGAYLISNLITKK 281
          |: : ||  :| ||  |:|  :|:  |  :        |:|||:|| | |:|| |:|
Db  633 KDTSSKLQQL-QLERANFEQKESTLSDENDLRTLKLLKLEESNKSLIKKQEDVDSL-EKN 690

Qy  282 ISAIKDSGELKAEIEKAKKCS----EEFTAKLKGEHTDLGKE-----GVTDDNAKKAIL 331
          |  :|:  |||::: ::|:|      | |::| |:| ||       ||:|:|||||:|
Db  691 IQTLKEDLRKSEEALFSKLEAKNLREVIDNLKGKHETLEAQRNDLHSSLSDAKNTNAIL 750

Qy  332 KTNNDKTKGADELEKLFESVKNL----SKAAKEMLTNSVKELTS 371
           |:|||||||||||||||||| |     ||||||:| ||||||||
Db  751 -SSELTKSSEDVKRLTANVETLTQDSKAMKQSFTSLVNSYQS 791
```

ID   RBP2_PLAVB   STANDARD;   PRT;   1251 AA.
AC   Q00799;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-APR-1993 (Rel. 25, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   RETICULOCYTE BINDING PROTEIN 2 (FRAGMENT).
GN   RBP2.
OS   Plasmodium vivax (strain Belem).
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=31273;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92315338; PubMed=1617731;
RA   Galinski M.R., Medina C.C., Ingravallo P., Barnwell J.W.;
RT   "A reticulocyte-binding protein complex of Plasmodium vivax
RT   merozoites.";
RL   Cell 69:1213-1226(1992).
CC   -!- FUNCTION: INVOLVED IN RETICULOCYTE ADHESION. SPECIFICALLY BINDS TO
CC   HUMAN RETICULOCYTE CELLS.
CC   -!- SUBCELLULAR LOCATION: MEMBRANE-BOUND (PROBABLE).

Right record

DR   M88098; AAA29744.1; -.
KW   Malaria; Receptor; Membrane.
FT   NON_TER   1
FT   NON_TER   1251
SQ   SEQUENCE   1251 AA; 143741 MW; 54BA51C7404AC572 CRC64;

Query Match   10.0%; Score 180.5; DB 1; Length 1251;
Best Local Similarity   21.8%; Pred. No. 0.3;
Matches 104; Conservative   63; Mismatches 160; Indels 151; Gaps 21;

```
Qy    5 NSGKDGNTS--ANSADESVKGPNLTEISKKITDSNAVLLAV------KEVEALLSSIDEL 56
          | |      : ||||:   |  :  |:|  |  ::|||||      |||||||||||:
Db  670 NEG--HGNVKQLVENIKELVDENNLSDILKQATGKNEEIQKIFHSTLNKAKTILGHVDTS 728

Qy   57 AKAIGKKI-----------------------KNDGSLDNE---ANRNE----SLLAGA 84
          || :|:|                        |||||||:    ||||:    ||||||
Db  729 AKVVGIKITPELALTELLGDAKILKTAQELKFESKNNVVLETENMSKNTNELDVHKNIQDA 788

Qy   85 YIISTLI----------TQKLSKLNGSEGL------KEKIAAAKKCSEEFSTKL- 122
          | | :|:           |||||||||||       |||:|||||||||||||
Db  789 YKVALEILAHSDEIDTRQKDSSKLIEMGNQIYLKVVLINQYKNKISSTKSKEEAVSVKIG 848

Qy  123 --KDNHAQLGIQGVTDENAKKAIL------------KANAGKDKGVEELEKLSG 163
            ||| |:| ||:   :|:||||            |||:||| ||||||:|:|
Db  849 NVSKKHSELSKITCSDKSYDNIIALEKQTELONLRNSFTQEKTNTNSDSK----LEKIKT 904

Qy  164 SLESLSRAAKEMLANSVKELTSPVVHGNNSGKDGNTSANSADESVKGPNLTEISKKITES 223
          ::||| :||||||||||||||| ||   :|  |:  ||: ||::     :|  |:|:
Db  905 DFESLKNALKTLEGE-----VNALKASSDNHEHVQSKSEPV-NPALSEIEKEET- 952

Qy  224 NAVVLAVKEVETLLTSIDELAK--------AIGKKIKNDVSLDNEADHNGSLI 268
          :|:  |::: | || :|| ||         ||| :||:| : |:|  |:  :
Db  953 ------DIDSLNTALDELLKKGRTCEVSRYKLIKDTVTKEISDDTELINTIEKNVK-- 1002

Qy  269 SGAYL-----------ISNLIT--KKISAIKDSGELKAEIEKAKKCSEEFTAKLKGEHTDL 316
          |||||           |||:|   |||||||||||||||||||||||||| |||||||||
Db 1003 --AYLAYIKKNYEDTVQDVLTLNEHFNTKQVSNHEPTNFDKSNKSSEELTKAVTDSKTII 1060

Qy  317 GKEGVTDDNAKKAILKTN-----NDKTKGADELEKLFESVKNLSKRAAKEM--LTNSVK 367
          |:|||:|||||||||||:     ||||||||||||||||||||:|:|||||  ||||||
Db 1061 SK------LKGVIIEVNENTEMNTIESSAKEIEALLYNELRNKKTISLNEIIYQTSNEVK 1111
```

RESULT 7
USO1_YEAST
ID   USO1_YEAST   STANDARD;   PRT;   1790 AA.
AC   P25386;
DT   01-MAY-1992 (Rel. 22, Created)
DT   01-MAY-1992 (Rel. 22, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   INTRACELLULAR PROTEIN TRANSPORT PROTEIN USO1.
GN   USO1 OR INTP1 OR YDL058W.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91185402; PubMed=2010462;
RA   Nakajima H., Hirata A., Ogawa Y., Yonehara T., Yoda K.,
RA   Yamasaki M.;
RT   "A cytoskeleton-related gene, uso1, is required for intracellular
RT   protein transport in Saccharomyces cerevisiae.";
RL   J. Cell Biol. 113:245-260(1991).
RN   [2]
RP   SEQUENCE OF 782-1790 FROM N.A.

RA   Hostetter M.K., Herman D.J., Bendel C.M., McClellan M., Tao N.,
RA   Kendrick K.E.;
RL   Submitted (FEB-1993) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 1-8 FROM N.A.
RA   Bai Y., Symington L.S.;
RL   Submitted (MAY-1996) to the EMBL/GenBank/DDBJ databases.
CC   COMPLEX.
CC   -!- FUNCTION: REQUIRED FOR PROTEIN TRANSPORT FROM THE ER TO THE GOLGI
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC. ASSOCIATED WITH INTRACELLULAR
CC   MEMBRANES. PROBABLY PRESENT ON VESICLES OPERATIONAL BETWEEN THE
CC   ER AND THE GOLGI COMPLEX.
CC   -!- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, COMPOSED
CC   OF AN HEPTAPEPTIDE REPEAT PATTERN CHARACTERISTIC OF ALPHA-HELICAL
CC   COILED COILS. MAY FORM FILAMENTOUS STRUCTURES IN THE CELL.
CC   -!- SIMILARITY: BELONGS TO THE VDP/USO1/YBL047C FAMILY.
CC   -!-

DR   EMBL; X54378; CAA38253.1; -.
DR   EMBL; L03188; AAB00143.1; -.
DR   EMBL; U53668; AAB66659.1; -.
DR   PIR; A38455; A38455.
DR   HSSP; P80220; 1DIP.
     SGD; S002216; USO1.
DR   InterPro; IPR002017; Spectrin.
KW   Transport; Protein transport; Cytoskeleton; Coiled coil.
FT   DOMAIN   1   724   GLOBULAR HEAD.
FT   DOMAIN   725   1790   COILED COIL (POTENTIAL).
FT   DOMAIN   465   487   CHARGED (HYPER-HYDROPHILIC).
FT   DOMAIN   991   1790   DISPENSABLE FOR THE PROTEIN FUNCTION.
FT   DOMAIN   1172   1786   ASP/GLU-RICH (ACIDIC).
FT   CONFLICT   847   847   G -> E (IN REF. 2).
FT   CONFLICT   924   924   E -> K (IN REF. 2).
FT   CONFLICT   1253   1253   V -> I (IN REF. 2).
FT   CONFLICT   1319   1319   V -> I (IN REF. 2).
FT   CONFLICT   1461   1461   N -> S (IN REF. 2).
FT   CONFLICT   1581   1581   L -> S (IN REF. 2).
FT   CONFLICT   1600   1600   I -> V (IN REF. 2).
FT   CONFLICT   1661   1661   R -> S (IN REF. 2).
FT   CONFLICT   1772   1772   D -> DEEDDEE (IN REF. 2).
SQ   SEQUENCE   1790 AA; 206424 MW; 6CE2B2165FD4818 CRC64;

Query Match
Best Local Similarity 23.7%;
Matches 89; Conservative 68; Mismatches 156; Indels 63; Gaps 13;

Score 178.5; DB 1; Length 1790;
Pred. No. 0.54;

```
Qy    44 KEVEALLSSIDELAKAIGKKI-KNDGSLDNEANRNESLLAGAYTISTLITQKLSKLN---   99
            |:|::|||  |  ||  ||||| | |:::| | | |  ||| ||||||||| |||:|
Db  1022 KNIEQLKKTISDLEQTKEEIISKSDSSKDEYESQOISLKEKLETATFANDENYNKISELT 1081

Qy   100 -GSEGLKEKIAAAKKCSEEFSTKLKDNHAQLGIQGVTDENAKKAILKANAAGKDKGVEEL  158
          |:::||  | |  ||||||::||||
Db  1082 KTREELEAELAAYKNLKNELETKLETSEKALKEVKENEEHLKEEKIQLEKEATET-KQQL 1140

Qy   159 EKLSGSLESLSK------AA-----KEMLANSVKELTSPVVHGNNSGKGNTSANSADESV  208
          |:|||| |:||      :AA
Db  1141 NSLRANLESLEKEHEDLAAQLKKYEEQIANKERQYNEEISQLN----DEITSTQQENESI 1196

Qy   209 KGPN-------------LTEISKKITESNAVVLAYKEV----ETLLTSIDELAKAIGKK  250
Db  1197 KKKNDELEGEVKAMKSTSEEQSNLKKSEIDALNLQIKELKKKNETNEASLLESIKSVSE 1256

Qy   251 IKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSG----ELKAEIEKAKKCSEEFT  306
Db  1257 TVKIKELQDECNFKEKEVSE---------LEDKLKASEDKNSKYLBLQKESEKIKEELDAKT 1309
```

```
Qy    10  GNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEAALLSSIDELAKAI--GKKIKND  67
              :::||:| |:|:|::||| |:::||||::::||::||::||:||||:|| |:|:| | |
Db   115  GKLGIDNADLKAR---ITELEKSVEEKNDVLSQIKK------ELEEAEKDIQFGREVH--  163

Qy    68  GSLDNEANRNESLLAGAYTISTLITQK---LSKLNGS-EGLKEKIAAAKKCGSEEFSTKLK  123
              |:||||| |:||:||||:|||||:||   |::|:||: |||:|||||||| ||:::|:|:
Db   164  -AADLLRHKQE-------IAEKENVISKLNGELQPLRQKVDETDRNLQQEKQKVL  210

Qy   124  DNHAQLGIGVTDENAKKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANSVKEL  183
              || :|||:||:|| |||:|| ||:||| :||||:|:|:||:|:||||||:|||||||
Db   211  SLEQQL---AVTKENAKKDFELA-ALGHOLADKEYNAKIAELES-------KIADKKDF  259

Qy   184  TSPVV---HGNN-------SGKDG------NTSANSADESVKGP--------------  211
              |::||   ||||       |||||       ||:|||||:|||||
Db   260  ELAALGHQHAHNEYQAKLAEKDGQIKQLEEQKQILDASRKGTTARDLEAVRQAKKATEAEL  319

Qy   212  -NLTEISKKITESNAVVLAVK----------EVETLLTSIDE---LAKAIGKKI  251
              :| ||||:|::|||:||||| |:           ||:||||:|   |||||||:
Db   320  NNLKAELAKVTEQKOILDASRKGTTARDLEAVRKSKKQQVEAALKQLEEQNKISEASRKGL  379

Qy   252  KNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAEIERAKKCGSEEFTAKLKG  311
              |:||:||||||:||:|||||:|||:|||| ||::|||:|:|| |:|:||||| ||:::|:|:
Db   380  ----KKQVEKDLANLTAELDKVKEEKQISDASRQG  420

Qy   312  EHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNL--------------SK  356
              |:| ||:||:||:|| |||:||:||:||||||:|:|:| |:||||             |||
Db   421  LRRDLDASREAKKQVEKA-LEEANSKLAALEKLNKDLEESKKLTEKEKAELQAKLREAAK  479

Qy   357  AAKEMLTNSVKEL  369
              ||:|||:||||:|
Db   480  ALKEQLAKQAEEL  492
```

```
Query Match             9.5%;  Score 171.5;  DB 1;  Length 1509;
Best Local Similarity   24.3%;
Matches   97;  Conservative   70;  Mismatches  143;  Indels   89;  Gaps   18;

Qy    19  ESVKGPNLTEISKKITDSNAVLLAVKEVEA---LLSSIDELAKAIGKKIKNDGSLDNEA  74
              |:|| |::|:||::|:||||||:||||||   |||:|||||| |:|:||||:|||||
Db   953  EAEKGELKASLEEERNRKALQEAKTKVESERNELQDKYEDEAAAHDSLKKEEDLSREL  1012

Qy    75  NRNESLLAGAYTISTLITQKLSKL-NGSEGLK---EKIAAAKKCGSEEFSTKLKDNHAQLG  130
              ||:|||||||:|||||:|:|:|| |||:|:|   |:|||||||||||::|:|||::||
Db  1013  RETKDALADAENISETLRSKLKNTERGADDVRNELDDVIATKLQLEKTKKSLEEELAQTR  1072

Qy   131  IQGVTDENAKKAILKANAAGKDKGVEELE------KLSGSLESLSKAAKEMLANS  179
              |::||||| | :||:|:| | |:|:||||      ||||:|:||||||||:||
Db  1073  AQLEEEKSGKEA---ASSKAKQLG-QQLEDARSEVDSLKSKLSAAEKSL-KTAKDQNRDL  1127

Qy   180  VKELTSPVVHGNNSGKDGNTSANSADESVKG--PNLTEISKKIT----ESNAVVLAVKEV  233
              |:|:||:| | |:| : ||:|||||||:|||  |::|::||::||    |||||||||||
Db  1128  DEQL-------EDERTVRANVDKQKKALEAKLTELEDQVTALDGQKNA---AAAQA  1173

Qy   234  ETLLTSIDEL--------AKAIGKKIKNDVSLLDNEADHNGSLISGAYLISNLITKKIS  283
              |:|||:|||          |:||||||| |:|| || |:||:||||||:|||:||::|
Db  1174  KTLKIQVDETKRRLEEAEASAARLEKERKN--ALDEVAQ-----------LTADLD  1216

Qy   284  AIKDSG------ELKAEIEKAKK---CSEEFTAKLKGEHTDLGKEGVTDDNARK  328
              |::|||      |||:|||:|:|   ||::|:|:||:||:|| ||:||:|||||
Db  1217  AERDSGAQQRRKLNTRISELQSELENAPKTGGASSEEVKRLEGELRERLEEELITAQEARA  1276

Qy   329  AILKTNNDKTKGADELEKLFESVKNLSKAAKEMLTNSVK  367
              |::|:|||:|:|||||||:|||||||:||||||:||:|
Db  1277  AAEK-NLDKAN--LELEELRQEADDAARDNDKLVKDNRK  1312
```

```
DT   01-OCT-1993 (Rel. 27, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   HYPOTHETICAL 84.1 KDA PROTEIN IN HEMY-GLTT INTERGENIC REGION (ORFB).
GN   YHGE.
OS   Bacillus subtilis.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Bacillus.
OX   NCBI_TaxID=1423;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168;
RA   Noback M.A., Terpstra P., Holsappel S., Venema G., Bron S.;
RL   Submitted (JUN-1997) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE OF 1-145 FROM N.A.
RX   MEDLINE=93094140; PubMed=1459957;
RA   Hansson M., Hederstedt L.;
RT   "Cloning and characterization of the Bacillus subtilis hemEHY gene
RT   cluster, which encodes protoheme IX biosynthetic enzymes.";
RL   J. Bacteriol. 174:8081-8093(1992).
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC   -!- SIMILARITY: TO L.LACTIS PHAGE INFECTION PROTEIN (PIP).
```
```
DR   EMBL; Y14083; CAA74522.1; -.
DR   EMBL; M97208; AAA22521.1; -.
DR   EMBL; Z99109; CAB12856.1; -.
DR   PIR; F47045; F47045.
DR   SubtiList; BG10433; yhgE.
KW   Hypothetical protein; Transmembrane; Complete proteome.
FT   TRANSMEM     16     36       POTENTIAL.
FT   TRANSMEM    585    605       POTENTIAL.
FT   TRANSMEM    625    645       POTENTIAL.
FT   TRANSMEM    655    675       POTENTIAL.
FT   TRANSMEM    742    762       POTENTIAL.
SQ   SEQUENCE   775 AA;  84093 MW;  49A5AAC8A6A70B66 CRC64;
```

```
Query Match            9.48;  Score  169.5;  DB 1;  Length 775;
Best Local Similarity  21.58%;  Pred. No. 0.53;
Matches   91;  Conservative   74;  Mismatches  141;  Indels  117;  Gaps  19;

Qy    5 NSGKD--GNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDELAKAI--   60
         :|||    | | :|::| ::||: ::|:||::::|:|::| :|||
Db  135 NAGSNYVGATIGEKAIDKLK----ASVSKEVTE------QYTKVIFDNFKDIAKGLSD    182

Qy   61 ---GKKIKNDGSLDNEANRNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEE  117
           || ||  || |  |  |: | ::: |: |  | |:  |: ||  |:| |:
Db  183 ASSGAKKIDDGTKDAK-----------NGSAQLKENLAKLKESTAT               217

Qy  118 FSTK---LKDNHAQL--GIQ----GVTDENAKKAILKAN--AAGKDKGVEELEKLSGS   164
          |||     |  ||:  |:|      :||  :| :|   ||   | ||:|  ||
Db  218 ISDKTAQLADGAAQVTSGIQSLDSSLGKFQDSSNQIYDKSSQLAAGSGELITSKMNELLAG 277

Qy  165 LESLSKAAKEMLANSVKELTSPVVHGNNSGGKDGNTSANSADESVKGPNLTEISKKITESN 224
          |:|:  |:| |  |:  | |   :    ||  :| | | |  |  |||  |:|: ||
Db  278 LQNVQKGTPN-LTNGLDQLNSKV----QEGSEKAARAEKIINALDLTKLETAVNNLE     329

Qy  225 AVVLAVKEVETLLTSIDELAKAIGKKKIKNDVSLDN--EADHNGSLISGA----------  271
          | |||  |  |  :|||| |: | | |  :  |    | | | |: |
Db  330 KSETAMKEFKQLTDFENSLKNRDQAFKNVINSSDFLTAEQKSQLINSVEKKLPQVDAPD   389

Qy  272 --YLISNLIT---KKISAIKDSGE-LKAEIERAKKCSEEFTAKILKGEHTDLGKEGVTDD  324
          |::|| |   :|: ||  |  :| || | ||   |:| ::| || |   | | ||
Db  390 FDQILSQLPTADQLPDIATIKSSLEDVKRAQVAQV-KAMPEATSKLY-----------   434

Qy  325 NAKKAILKTNNDKTKGADEL------EKLFESVKNL--------SKAAKEMLTNSVK    367
          |||||| | || |  | |      || || |:|        |  | ||| | |
Db  435 NGAKTIQDAIDRLITEGADKIYNGSQKLTDGQTKLTAGIGETNKQFRAKAKAGSEQLVTGSS 494

Qy  368 ELT    370
         :::
Db  495 QVS    497
```

```
RESULT 11
MYSC_CAEEL
ID   MYSC_CAEEL     STANDARD;      PRT;   1947 AA.
AC   P12845;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   MYOSIN HEAVY CHAIN C (MHC C).
GN   MYO-2.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RX   MEDLINE=85201409; PubMed=3888374;
RA   Karn J., Dibb N.J., Miller D.M.;
RT   "Cloning nematode myosin genes.";
RL   Cell Muscle Motil. 6:185-237(1985).
RN   [2]
RP   SEQUENCE OF 1-21; 112-371 AND 1501-1772 FROM N.A.
RC   STRAIN=Bristol N2;
RX   MEDLINE=89178677; PubMed=2926820;
RA   Dibb N.J., Maruyama I.N., Krause M., Karn J.;
RT   "Sequence analysis of the complete Caenorhabditis elegans myosin
RT   heavy chain gene family.";
RL   J. Mol. Biol. 205:603-613(1989).
CC   -!- FUNCTION: MUSCLE CONTRACTION.
CC   -!- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
CC       HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
CC       AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
CC   -!- SUBCELLULAR LOCATION: THICK FILAMENTS OF THE MYOFIBRILS.
CC   -!- TISSUE SPECIFICITY: FOUND EXCLUSIVELY IN THE PHARYNGEAL MUSCLE.
CC   -!- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CC       CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CC       CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC   -!- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
CC       ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
CC   -!- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
CC       MEROMYOSIN (LMM) AND 1 HEAVY MEROMYOSIN (HMM). IT CAN LATER BE
CC       SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
CC       SUBFRAGMENT (S2).
CC   -!- MISCELLANEOUS: THERE ARE FOUR DIFFERENT MYOSIN HEAVY CHAINS IN C.
CC       ELEGANS.
CC   -!- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
```
```
DR   EMBL; X08066; CAA30855.1; -.
DR   EMBL; M37233; AAA28121.1; -.
DR   EMBL; M37235; AAA28122.1; -.
DR   EMBL; M37236; AAA28123.1; -.
DR   PIR; S05697; S05697.
DR   HSSP; P08799; 1MND.
DR   InterPro; IPR002928; Myosin_tail.
DR   InterPro; IPR001609; myosin_head.
```

RESULT 12
MYH8_HUMAN

ID   MYH8_HUMAN   STANDARD;   PRT;   1937 AA.
AC   P13535; Q14910;
DT   01-JAN-1990 (Rel. 13, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   MYOSIN HEAVY CHAIN, SKELETAL MUSCLE, PERINATAL (MYHC-PERINATAL).
GN   MYH8.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90325631; PubMed=2373371;
RA   Karsch-Mizrachi I., Feghali R., Shows T.B. Jr., Leinwand L.A.;
RT   "Generation of a full-length human perinatal myosin heavy-chain-
RT   encoding cDNA.";
RL   Gene 89:289-294(1990).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Skeletal muscle;
RX   MEDLINE=95324556; PubMed=7601129;
RA   Jullian E.H., Kelly A.M., Pompidou A.J., Hoffman R., Schiaffino S.,
RA   Stedman H.H., Rubinstein N.A.;
RT   "Characterization of a human perinatal myosin heavy-chain
RT   transcript.";
RL   Eur. J. Biochem. 230:1001-1006(1995).
RN   [3]
RP   SEQUENCE OF 502-1937 FROM N.A.
RC   TISSUE=Skeletal muscle;
RX   MEDLINE=90235862; PubMed=1691980;
RA   Bober E., Buchberger-Seidl A., Braun T., Singh S., Goedde H.W.,
RA   Arnold H.H.;
RT   "Identification of three developmentally controlled isoforms of human
RT   myosin heavy chains.";
RL   Eur. J. Biochem. 189:55-65(1990).
RN   [4]
RP   SEQUENCE OF 860-1937 FROM N.A.
RX   MEDLINE=89234168; PubMed=2715179;
RA   Feghali R., Leinwand L.A.;
RT   "Molecular genetic characterization of a developmentally regulated
RT   human perinatal myosin heavy chain.";
RL   J. Cell Biol. 108:1791-1797(1989).
RN   [5]
RP   SEQUENCE OF 1-46 FROM N.A.
RA   Esser K., Tidhar A., Myszkowski M.;
RT   "Isolation and characterization of the human perinatal MHC promoter.";
RL   Submitted (MAY-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: MUSCLE CONTRACTION.
CC   -!- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
CC   HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
CC   AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
CC   -!- SUBCELLULAR LOCATION: THICK FILAMENTS OF THE MYOFIBRILS.
CC   -!- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CC   CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CC   CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC   -!- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
CC   ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
CC   -!- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
CC   MEROMYOSIN (LMM) AND 1 HEAVY MEROMYOSIN (HMM). IT CAN LATER BE
CC   SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
CC   SUBFRAGMENT (S2).
CC   -!- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 IQ DOMAIN.

DR   EMBL; M36769; AAC17185.1; -.
DR   EMBL; Z38133; CAA86293.1; -.
DR   EMBL; X51592; CAA35941.1; -.
DR   EMBL; M35250; AAA36346.1; -.
DR   EMBL; AF067143; AAC21557.1; -.
DR   PIR; A30220; A30220.
DR   HSSP; P13538; 2MYS.
DR   MIM; 160741; -.
DR   InterPro; IPR000048; IQ.
DR   InterPro; IPR000928; Myosin_tail.
DR   InterPro; IPR002928; Myosin_tail.
DR   InterPro; IPR001609; Myosin_head.
DR   Pfam; PF00612; IQ; 1.
DR   Pfam; PF00063; Myosin_head; 1.
DR   Pfam; PF01576; Myosin_tail; 1.
DR   PRINTS; PR00193; MYOSINHEAVY.
DR   ProDom; PD000355; myosin_head; 1.
DR   SMART; SM00242; MYSc; 1.
KW   Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
KW   ATP-binding; Methylation; Multigene family.
FT   DOMAIN   1   855   MYOSIN HEAD-LIKE.
FT   DOMAIN   856   1947   RODLIKE TAIL (S2 AND LMM DOMAINS).
FT   DOMAIN   856   1947   COILED COIL (POTENTIAL).
FT   NP_BIND   174   181   ATP (BY SIMILARITY).
FT   DOMAIN   668   690   ACTIN-BINDING.
FT   DOMAIN   773   787   ACTIN-BINDING.
FT   MOD_RES   125   125   METHYLATION (TRI-) (POTENTIAL).
FT   MOD_RES   708   708   ALKYLATION (SH-1).
FT   MOD_RES   718   718   ALKYLATION (SH-2).
FT   CONFLICT   132   132   E -> D (IN REF. 2).
FT   CONFLICT   137   137   M -> I (IN REF. 2).
SEQUENCE   1947 AA; 223009 MW; 6D65AE99BD013627 CRC64;

Query Match   9.4%; Score 169.5; DB 1; Length 1947;
Best Local Similarity   22.9%; Pred. No. 1.5;
Matches   93; Conservative   61; Mismatches 118; Indels 135; Gaps 18;

Qy   23 GPNLTEISKKITDSNAVLLAVKEVEALLSSIDELAKAIGKKIKNDG----------SLD 71
              |:||:||| || |:|| ::: |:|:|: ||
Db   1202 GEQLTVLKKGSDA------------IQELSDQIEQLOKGRIEKEKGHMQREFEDESCAALD 1252

Qy   72 NEA--NRNESLLAGAYTIST----------------LITQKLSKLNGS-------EGLK 105
              ::| ||||| | | | | |
Db   1253 QEAKLRADQERIAKGYEVQTSELRLRKADEQSRQLQDFVSSK-GRLNSENSDLARQVEELE 1311

Qy   106 EKIAAAKKCSEEFSTKLKDNHAQLGIQGVTDENAKKAILKANAAGKDKGVEELEKLSGSL 165
              |:|||:|::| ::| |:| :|||| ||: |:| | || | | | || :
Db   1312 AKIQAANRLKLQFSNEL--DHAKRQAEEEASRERQNLSNLSKNLA------RELEQLKESI 1363

Qy   166 E--------------SLSKAAKEMLANSVKELTSPVVHGNNSGKDGNTSANSADESVKGPN-- 212
              | |||| | | | |
Db   1364 EDEVAGKNEASRQLSKASVELDQWRTKFET-----------------EGLIGADEFDEVKKRQNQK 1412

Qy   213 LTEISKKITESNAVVLAVKEVETILLTISIDELAKAIGKKIKNDVSLDNEADHINGSLISGAY 272
Db   1413 TSEIQDALDACNAKIVALENARSRLT----AEADANRL------EAEHAQAVSS-- 1457

Qy   273 LISNLITKKISAI------KDSGELKAEIEKAKKCSEEFTA---KLKGEHTDLGKEGV 321
Db   1458 ----LEKKQKAFDKVIDEWKKKKVDDLYLELDGAQRDARQLSGEAHKLRGQHDTL----- 1507

Qy   322 TDDNAKKAILKTNNDKTKGADELEKLFESVKNLSRAAKEMLTNSVKE 368
Db   1508 ----------------ADQVEGLRRENKSLSDETRD-LTESLSE 1534

```
DR   SMART; SM00015; IQ; 1.
DR   SMART; SM00242; MYSc; 1.
DR   PROSITE; PS50096; IQ; 1.
KW   Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
KW   ATP-binding; Methylation; Alkylation; Multigene family;
KW   Calmodulin-binding;
FT   DOMAIN        1    780       MYOSIN HEAD-LIKE.
FT   DOMAIN      781    813       IQ.
FT   DOMAIN      842   1937       COILED COIL (POTENTIAL).
FT   NP_BIND     181    188       ATP.
FT   DOMAIN      [2]                ACTIN-BINDING.
FT   DOMAIN      658    680       ACTIN-BINDING.
FT   DOMAIN      760    774       ACTIN-BINDING.
FT   MOD_RES     132    132       METHYLATION (TRI-) (POTENTIAL).
FT   MOD_RES     698    698       METHYLATION (POTENTIAL).
FT   MOD_RES     708    708       ALKYLATION (SH-1) (POTENTIAL).
FT   CONFLICT     15     15       ALKYLATION (SH-2) (POTENTIAL).
FT   CONFLICT    970    970       A -> R (IN REF. 2).
FT   CONFLICT   1072   1072       E -> Q (IN REF. 1 AND 4).
FT   CONFLICT   1247   1247       M -> N (IN REF. 3).
FT   CONFLICT   1251   1252       N -> D (IN REF. 1 AND 4).
FT   CONFLICT   1261   1261       MC -> DGG (IN REF. 3).
FT   CONFLICT   1261   1261       E -> G (IN REF. 1 AND 4).
FT   CONFLICT   1297   1297       K -> Q (IN REF. 1 AND 4).
FT   CONFLICT   1377   1378       KY -> NT (IN REF. 3).
FT   CONFLICT   1504   1505       EN -> KY (IN REF. 1 AND 4).
FT   CONFLICT   1504   1505       EN -> AH (IN REF. 3).
FT   CONFLICT   1847   1847       D -> Y (IN REF. 1 AND 4).
FT   CONFLICT   1914   1914       D -> H (IN REF. 2).
SQ   SEQUENCE 1937 AA;  222762 MW;  A3EE2D151792E9E8 CRC64;
```

Query Match   9.3%;   Score 168;   DB 1;   Length 1937;
Best Local Similarity   23.2%;   Pred. No. 1.7;
Matches   88;   Conservative   70;   Mismatches 166;   Indels 56;   Gaps 14;

```
Qy    15  NSADESVKGPNLTEISKKITDSNAVLLAVK------EVEALLSSIDELAKAIGKKIKNDGS   69
              :::            :: : ::           : ::     ::     ::: :
Db   914  NKIQLEAKIKEVTERAEEEEINAELTAKKRKLEDECCSELKKDIDDLELTLAKVEKEKHA  973

Qy    70  LDNEANRNESLLAGAYTISTLLITQKLSKLNGSEGLKEKIAAAKKCSEEFSTKLKDNHAQL  129
              : :: :          ::    :  :     ::  :    :   :    :  :  ::
Db   974  TENKVKN---------------GLDETIA----KLSKE-KKALQETHQQT 1009

Qy   130  GIQGVTDENAKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANSVKELTSPVVH  189
              :    : :::: :   : :: ::  :::  : :: :   :   : :   :  :::
Db  1010  LDDLQAEEDKVNILTKA-----KTKLEQGQVDDLEGSLEQEKKLRMD-LERAKRKLEGDLKL 1064

Qy   190  GNNSGKDGNTSANSADESV--KGPNLTEISKKITESNAVVLAV-KEVETLLTSIDEL---  243
              :: :     :::        :    ::::  ::::  :::      :  :::
Db  1065  AQESTMDMENDKQQLDEKLEKKEFEISNLISKIEDEQAVEIQLQKKIKELQARIEELGEE 1124

Qy   244  ------AKAIGKKIKNDVSLD-NEADHNGSLISGAYLISNLITPKKISAIKDSGELKAEIE  296
                     ::    : ::  :    :::    : :   ::::   :  :::  ::  :
Db  1125  IEAERASRAKAEKQRSDLSRELEEISERLEEAGGATSAQVELNKKREA--EFQKLRRDLE 1182

Qy   297  KAKKCSEEFTAKLKGEHPD-LGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLS  355
              ::  :: :::        ::         : :: : :  :   : :::      ::
Db  1183  EATLQHEAMVAALRKKHADSMAELGEQIDNLQRVKQKLEKEKSELKMETDDLSSNAEATS 1242

Qy   356  KAAKEM------LTNSVKEL  369
              :::           ::: :
Db  1243  KAKGNLEKMCRSLEDQVSEL 1262
```

RESULT 13
ID   MYSB_CAEEL
```
AC   P02566;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   MYOSIN HEAVY CHAIN B (MHC B).
GN   UNC-54 OR MYO-4.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
```

(second column)

```
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=83273600; PubMed=6576334;
RA   Karn J., Brenner S., Barnett L.;
RT   "Protein structural domains in the Caenorhabditis elegans unc-54
RT   myosin heavy chain gene are not separated by introns.";
RL   Proc. Natl. Acad. Sci. U.S.A. 80:4253-4257(1983).
RN   [2]
RP   SEQUENCE OF 850-1966 FROM N.A.
RX   MEDLINE=82272395; PubMed=202124;
RA   McLachlan A.D., Karn J.;
RT   "Periodic charge distributions in the myosin rod amino acid sequence
RT   match cross-bridge spacings in muscle.";
RL   Nature 299:226-231(1982).
RN   [3]
RP   SEQUENCE OF 1876-1966 FROM N.A.
RX   MEDLINE=83232892; PubMed=6571695;
RA   Wills N., Gesteland R.F., Karn J., Barnett L., Bolten S.,
RA   Waterston R.H.;
RT   "The genes sup-7 X and sup-5 III of C. elegans suppress amber
RT   nonsense mutations via altered transfer RNA.";
RL   Cell 33:575-583(1983).
CC   -!- FUNCTION: MUSCLE CONTRACTION.
CC   -!- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
CC       HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
CC       AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
CC   -!- SUBCELLULAR LOCATION: THICK FILAMENTS OF THE MYOFIBRILS.
CC   -!- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CC       CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CC       CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC   -!- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
CC       ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
CC   -!- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
CC       MEROMYOSIN (LMM) AND 1 HEAVY MEROMYOSIN (HMM). IT CAN LATER BE
CC       SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
CC       SUBFRAGMENT (S2).
CC   -!- MISCELLANEOUS: THERE ARE FOUR DIFFERENT MYOSIN HEAVY CHAINS IN C.
CC       ELEGANS.
CC   -!- MISCELLANEOUS: MHC A AND MHC B ARE FOUND EXCLUSIVELY IN THE BODY
CC       WALL MUSCLE. THEY CO-ASSEMBLE INTO BODY WALL THICK FILAMENT.
CC   -!- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; J01050; AAA28124.1; -.
DR   EMBL; V01494; CAA24738.1; -.
DR   PIR; A02922; MWKW.
DR   HSSP; P08799; 1MND.
DR   InterPro; IPR002928; Myosin_tail.
DR   InterPro; IPR001609; myosin_head.
DR   Pfam; PF00063; myosin_head; 1.
DR   Pfam; PF01576; myosin_tail; 1.
DR   PRINTS; PR00193; MYOSINHEAVY.
DR   ProDom; PD000355; myosin_head; 1.
DR   SMART; SM00242; MYSc; 1.
KW   Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
KW   ATP-binding; Methylation; Alkylation; Multigene family;
FT   DOMAIN        1    850       MYOSIN HEAD-LIKE.
FT   DOMAIN      851   1966       COILED COIL (POTENTIAL).
FT   DOMAIN      851   1164       ALPHA-HELICAL TAILPIECE (S2).
FT   DOMAIN     1165   1176       HINGE.
FT   DOMAIN     1165   1966       LIGHT MEROMYOSIN (LMM).
FT   NP_BIND     177    184       ATP (BY SIMILARITY).
FT   DOMAIN      665    687       ACTIN-BINDING.
FT   DOMAIN      769    783       ACTIN-BINDING.
```

```
FT   MOD_RES      128    128    METHYLATION (TRI-) (POTENTIAL).
FT   MOD_RES      705    705    ALKYLATION (SH-1).
FT   MOD_RES      715    715    ALKYLATION (SH-2).
FT   CONFLICT    1337   1337    E -> V (IN REF. 2).
FT   CONFLICT    1880   1880    I -> R (IN REF. 2).
FT   CONFLICT    1880   1880    L -> L (IN REF. 2).
SQ   SEQUENCE    1966 AA;  225125 MW;  B66F0BB2FE27B67F CRC64;
```

Query Match   97; Conservative   79; Mismatches   167; Indels   82; Gaps   19;
Best Local Similarity   22.8%; Score 167; DB 1; Length 1966;
Matches   9.2%; Pred. No. 1.9;

```
Qy      8  KDGNTSANSADESVKGP-----------------------NLTEISKKITDSNAVLLAVKEVEALLS   51
           ||||||| | | |:    |:|:                :|||:|| |||:|||:|||:|||||:
Db   1121  KDGQSRISELEELENEROSRSKADRAKSDLQORELEELGEKLDEQGGATAAQVEVNKRE  1180

Qy     52  SIDELAKAIGKKIKNDGSLDNEANRN-ESLLAGAYTIST----LITOKLSKLNGSEGLKE  106
           |:|:|        ||| |:    | |   ::|            |:| ||  | |||
Db   1181  A--ELA---KLRRD---LEEANMHENOLGGLRKKHTDAVAELTDOLDQLNKARAKVE  1230

Qy    107  KIAA-----------------------IOGVTDENAKKAILKANA  148
           || |                         |::|:|||| ||| |
Db   1231  KDKAQAVRDAEDLAAOLDQEFTSGKL--NNEKLAKOFELQLTELQSKADEQSRQLDFTSL  1288

Qy    149  AGK-----------DKGVEELEKLSGSLESLSKAAKEMLANSVKELTSPVVIGNNS  193
           |||           :| |||||||:||:||| ||| || ||::|:|| |:
Db   1289  KGRLHSENGDLVRQLEDAESQVNOLTRIKSQITSOLEEARTADEEAREROTVAAQAKNY  1348

Qy    194  GKDGNTSANSADESVKGPNLIEISKKITESNAVVLAVK---EVETLLTSIDELAKRAIGKK  250
           |||   | | | |::|| :|:||| |:|| |||| |     :|||:|:| ||||  | |
Db   1349  QHEAEQLOESLEEEIGCKN--EILROLSKANADIOQWKARFEGEGLLKA-DELEDAKRRQ  1405

Qy    251  IK---NDV--SLDNEADHNGSL-ISGAYLISNLITKKISAIKDSGELKAEIEKAKCCSEEF  305
           ||   |:|    |||||| |:|  |: | |  | :| |::|:| | |:|| | | | |:||
Db   1406  AOKINELOEALDAANSKNASLEKTKSRLVGDLDDAQVDVERANGVASA-LEKKQKGFDKI  1464

Qy    306  TAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLSKAAKEMLTNS  365
           |||:|  | : | ||:||| ||| |||::| ||:   | || ::|:|||:||||:|::|
Db   1465  IDEWRKKTDDLAAE---LDGAQRDLRNTSTDLFKAKNAQEELAEVVEGLIRRENKS-LSQE  1520

Qy    366  VKELT  370
           :|:||
Db   1521  IKDLT  1525
```

```
RESULT 14
LMA3_HUMAN    STANDARD;    PRT;    1713 AA.

AC   Q16787; Q13679; Q13680;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   LAMININ ALPHA-3 CHAIN PRECURSOR (EPILIGRIN 170 KDA SUBUNIT) (E170).
GN   LAMA3.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Keratinocytes;
RX   MEDLINE=94357926; PubMed=8077230;
RA   Ryan M.C., Tizard R., Vandevanter D.R., Carter W.G.;
RT   "Cloning of the LamA3 gene encoding the alpha 3 chain of the adhesive
RT   ligand epiligrin. Expression in wound repair.";
RL   J. Biol. Chem. 269:22779-22787(1994).
RN   [2]
RP   SEQUENCE OF 1-1331 FROM N.A. (ISOFORMS A AND B).
RX   MEDLINE=96163880; PubMed=8586427;
RA   Vidal F., Baudoin C., Miquel C., Galliano M.-F., Christiano A.M.,
RA   Uitto J., Ortonne J.-P., Meneguzzi G.;
RT   "Cloning of the laminin alpha 3 chain gene (LaMA3) and identification
RT   of a homozygous deletion in a patient with Herlitz junctional
```

```
RT   epidermolysis bullosa.";
RL   Genomics 30:273-280(1995).
RN   [3]
-!- FUNCTION: BINDING TO CELLS VIA A HIGH AFFINITY RECEPTOR, LAMININ
    IS THOUGHT TO MEDIATE THE ATTACHMENT, MIGRATION, & ORGANIZATION OF
    CELLS INTO TISSUES DURING EMBRYONIC DEVELOPMENT BY INTERACTING
    WITH OTHER EXTRACELLULAR MATRIX COMPONENTS.
-!- FUNCTION: LAMININ-5 IS THOUGHT TO BE INVOLVED IN (1) CELL ADHESION
    VIA INTEGRIN ALPHA-3/BETA-1 IN FOCAL ADHESION AND INTEGRIN ALPHA-
    6/BETA-4 IN HEMIDESMOSOMES, (2) SIGNAL TRANSDUCTION VIA TYROSINE
    PHOSPHORYLATION OF PP125-FAK AND P80, (3) DIFFERENTIATION OF
    KERATINOCYTES.
-!- SUBUNIT: LAMININ IS A COMPLEX GLYCOPROTEIN, CONSISTING OF THREE
    DIFFERENT POLYPEPTIDE CHAINS (ALPHA, BETA, GAMMA), WHICH ARE BOUND
    TO EACH OTHER BY DISULFIDE BONDS INTO A CROSS-SHAPED MOLECULE
    COMPRISING ONE LONG & THREE SHORT ARMS WITH GLOBULES AT EACH END.
    THE ALPHA-3 CHAIN IS A SUBUNIT OF LAMININ-5 (EPILIGRIN/KALININ/
    NICEIN), AND POSSIBLY ALSO A COMPONENT OF LAMININ-6 (K-LAMININ)
    AND LAMININ-7 (KS-LAMININ).
-!- SUBCELLULAR LOCATION: EXTRACELLULAR; FOUND IN THE BASEMENT
    MEMBRANES (MAJOR COMPONENT).
-!- ALTERNATIVE PRODUCTS: THE TWO ISOFORMS A AND B DIFFER IN THEIR N-
    TERMINUS. THE SEQUENCE SHOWN HERE IS THAT OF THE SMALLER VARIANT
    A.
-!- TISSUE SPECIFICITY: SKIN; RESPIRATORY, URINARY, AND DIGESTIVE
    EPITHELIA AND IN OTHER SPECIALIZED TISSUES WITH PROMINENT
    SECRETORY OR PROTECTIVE FUNCTIONS. EPITHELIAL BASEMENT MEMBRANE,
    AND EPITHELIAL CELL TONGUE THAT MIGRATES INTO A WOUND BED. A
    DIFFERENTIAL AND FOCAL EXPRESSION OF THE ALPHA-3 CHAIN IS OBSERVED
    IN THE CNS.
-!- INDUCTION: LAMININ-5 IS UP-REGULATED IN WOUND SITES OF HUMAN SKIN.
-!- DOMAIN: THE ALPHA-HELICAL DOMAINS I AND II ARE THOUGHT TO INTERACT
    WITH OTHER LAMININ CHAINS TO FORM A COILED COIL STRUCTURE.
-!- DOMAIN: DOMAIN G IS GLOBULAR.
-!- DISEASE: DEFECTS IN LAMA3 ARE A CAUSE OF JUNCTIONAL EPIDERMOLYSIS
    BULLOSA (JEB) GRAVIS, ALSO KNOWN AS HERLITZ TYPE JEB. IT IS A
    BLISTERING DISORDER IN SKIN THAT IS CHARACTERIZED BY A SEPARATION
    OF BASAL CELLS FROM THE BASEMENT MEMBRANE DUE TO A DECREASED
    NUMBER OF HEMIDESMOSOMES. LAMININ-5 IS MISSING FROM THE BASEMENT
    MEMBRANE OF PATIENTS WITH THE GRAVIS FORM OF EPIDERMOLYSIS
    BULLOSA.
-!- SIMILARITY: CONTAINS 2.5 LAMININ EGF-LIKE DOMAINS.
-!- SIMILARITY: CONTAINS 5 LAMININ G-LIKE DOMAINS.
```

```
DR   EMBL; L34155; AAA59483.1; -.
DR   EMBL; X85107; CAA59428.1; -.
DR   EMBL; X85108; CAA59429.1; -.
DR   HSSP; P02468; 1TLE.
DR   MIM; 226700; -.
DR   MIM; 600805; -.
DR   InterPro; IPR000561; EGF-like.
DR   InterPro; IPR002049; Laminin_EGF.
DR   InterPro; IPR001791; Laminin_G.
DR   InterPro; IPR003129; TSPN.
DR   Pfam; PF00053; laminin_EGF; 2.
DR   Pfam; PF00054; laminin_G; 2.
DR   SMART; SM00180; EGF_Lam; 2.
DR   SMART; SM00282; LamG; 5.
DR   SMART; SM00210; TSPN; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS00186; EGF_2; 1.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS01248; LAMININ_TYPE_EGF; 2.
KW   Glycoprotein; Basement membrane; Extracellular matrix; Coiled coil;
KW   Laminin EGF-like domain; Cell adhesion; Repeat; Signal;
KW   Alternative splicing; Epidermolysis bullosa.
SIGNAL          1      20       POTENTIAL.
```

```
FT   CHAIN        21   1713   LAMININ ALPHA-3 CHAIN.
FT   DOMAIN       46    201   DOMAIN III A.
FT   DOMAIN       67    185   2.5 X LAMININ EGF-LIKE REPEATS.
FT   DOMAIN       67    113   LAMININ EGF-LIKE 1.
FT   DOMAIN      114    166   LAMININ EGF-LIKE 2.
FT   DOMAIN      167    185   LAMININ EGF-LIKE 3 (INCOMPLETE).
FT   DOMAIN      202    793   DOMAIN II AND I (HEPTAT REPEATS).
FT   DOMAIN      794   1713   5 X LAMININ G-LIKE REPEATS (DOMAIN G).
FT   DOMAIN      794    970   LAMININ G-LIKE 1.
FT   DOMAIN      971   1139   LAMININ G-LIKE 2.
FT   DOMAIN     1140   1353   LAMININ G-LIKE 3.
FT   DOMAIN     1354   1529   LAMININ G-LIKE 4.
FT   DOMAIN     1530   1713   LAMININ G-LIKE 5.
FT   DISULFID    231    327   BY SIMILARITY.
FT   DISULFID    396    548   BY SIMILARITY.
FT   DISULFID    594    621   BY SIMILARITY.
FT   DISULFID    702    765   BY SIMILARITY.
FT   DISULFID   1686   1713   BY SIMILARITY.
FT   DISULFID     67     76   BY SIMILARITY.
FT   DISULFID     69     83   BY SIMILARITY.
FT   DISULFID     86     95   BY SIMILARITY.
FT   DISULFID     98    111   BY SIMILARITY.
FT   DISULFID    114    126   BY SIMILARITY.
FT   DISULFID    116    135   BY SIMILARITY.
FT   DISULFID    137    146   BY SIMILARITY.
FT   DISULFID    149    164   BY SIMILARITY.
FT   DISULFID    202    202   INTERCHAIN (PROBABLE).
FT   DISULFID    205    205   INTERCHAIN (PROBABLE).
FT   SITE        658    660   CELL ATTACHMENT SITE (POTENTIAL).
FT   CARBOHYD    542    542   N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    645    645   N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    745    745   N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    882    882   N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    964    964   N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD   1108   1108   N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD   1131   1131   N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD   1325   1325   N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD   1477   1477   N-LINKED (GLCNAC...) (POTENTIAL).
FT   VARSPLIC      1     45   ... (IN ISOFORM B).
FT   CONFLICT      5      5   W -> R (IN REF. 2).
FT   CONFLICT    123    125   ATG -> GMC (IN REF. 2).
FT   CONFLICT    481    481   M -> K (IN REF. 2).
FT   CONFLICT    754    754   R -> L (IN REF. 2).
FT   CONFLICT    969    969   E -> Q (IN REF. 2).
FT   CONFLICT   1052   1052   D -> A (IN REF. 2).
FT   CONFLICT   1184   1184   G -> A (IN REF. 2).
SQ   SEQUENCE   1713 AA;  189304 MW;  45EA9BE1017B60D3 CRC64;

Query Match           9.2%;  Score 166.5;  DB 1;  Length 1713;
Best Local Similarity  23.0%;  Pred. No. 1.7;
Matches   90;  Conservative  71;  Mismatches 138;  Indels  93;  Gaps  16;

Qy   9  DGNTSANSADESVKGPNLTEISKKITDSNAVL----LAVKEVEALLSSIDELAKAIGKKI  64
            :  |||  ||||:  |  ||  |:|   :| :  ::
Db 396  ENNGLANSIRDS-----LNEYEAKLSDLRARLQEAAAQAKQANGLNQENERALGAIORQV 450

Qy  65  KNDGSLDNEAANRNESLLAGAYITSTLLITQKLSKLNGSEGLKEKIAAA-KKCSEEFSTKLK 123
Db 451  KEINSLQSDFTKYLTTADSSLLQTNIALQLMEK---SQKEYERLAASLNEARQELSDKVR 507

Qy 124  DNHAQLGIQGVTDENAKKAILKANAAGKDKGVEELEKLSGSLESLSRAAKEMLAN-SVKE 182
Db 508  E--...-LSRSACKTSLVEEAEKHARSLQELAKAQLEEIKRNASGDE 547

Qy 183  L:T-------...-SPVVHGNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKE 232
Db 548  LVRCAVDAATAYENILNAIKAAEDAANRAASASES----...----ALQTVIKE 590
```

```
DR   EMBL; M88097; AAA29743.1; -.
DR   HSSP; P36956; 1AM9.
KW   Malaria; Receptor; Signal; Transmembrane.
FT   SIGNAL        1     17   POTENTIAL.
FT   CHAIN        18   2869   RETICULOCYTE BINDING PROTEIN 1.
FT   DOMAIN       18   2807   EXTRACELLULAR.
FT   TRANSMEM   2808   2826   POTENTIAL.
FT   DOMAIN     2827   2869   CYTOPLASMIC.
FT   SITE       1030   1032   CELL ATTACHMENT SITE (POTENTIAL).
FT   SITE       2599   2601   CELL ATTACHMENT SITE (POTENTIAL).
SQ   SEQUENCE   2869 AA;  330213 MW;  B9DBE44220SEBCFF CRC64;

Qy 233  -----VETLLTSIDEL---AKAIGKKIKNDVS-LDN-EADHNGSLISGAYLISNLITKK 281
Db 591  DLPRKAKTLSSSNSDKLLNEAKMTQKKLKQEVSPALNNLQQTLNIVTVQKEVIDTNLTTLR 650

Qy 282  ---------ISAIRDSGELKAEIEKAKKCSEEFTAKLLKGEHTDLGKEGVTD------D 324
Db 651  DGLHGTQRGDIDAMISSA--KSMVRRANDITFDEVLDGLNPIQTDV--ERIKDTYGRTQNE 706

Qy 325  NAKKAILKTNNDKTKGADELEKLFESVKNLSK 356
Db 707  DFKKALTDADNSVNKLITNKLPDLWRKIESINQ 738

Query Match           9.18%;  Score 165.5;  DB 1;  Length 2869;
Best Local Similarity  22.48%;  Pred. No. 3.4;
Matches  101;  Conservative  63;  Mismatches 124;  Indels 163;  Gaps  20;

Qy  25  NLTEISKKIT-DSNAVLLAVKEVEALLSSIDELAKAIGKKIKNDGSLDNEANRNESLLAG  83
Db 1141 HINTVAHGITSNKNEILKSVKEVEDRKLNLVEQ--NEDYKKVKNP--ENE------- 1185

Qy  84  AYTISTLLITQKLSKLNGS-EGLKEKIAAAKCSEEFSTKLKDNHAQLGIQGVTDENAKKA 142
Db 1186 -------------KOLEAIRGSMSKLKEVI------...-NKH----VSEMTQLESTAN 1217

Qy 143  ILKANAAGKDK--GVEELEKLSGSL----ESLSKAAKEMLANSVKELTSPVVHGNNSGKD 196
```

```
Db 1218 TLKSNAKGKENEHDLEELNKTKGQMRDIYEKLKKIAEELKEGTVNEL--------------KD 1266
            ||                   | :|   |
Qy  197 GNTSANSADESVKGPNLTEISKKIT-ESNAVVLAVKEVETLLTSIDELAK------------ 245
           ||  : ::|| ::||  | :||||   :
Db 1267 ANEKANKVEPEPERNIIGHVLERITVEKDKAGKVVEEMNSLKTKIEKLIQETSDDSQNEL 1326

Qy  246 ---AIGKKIKN------DVSLDNEAD------------------------HNGSLISGA 271
           :| | |:|    | | |                         :
Db 1327 VTTSITKHLENAKGYEDVIKRNEEDSIQLREKAKSLETLDEMKKLVQQVNMNLQSAIQGN 1386

Qy  272 YLISN----------LITKKISAI---------------------KDSGELK---AE 294
        ||             |:  |:|                        | | |     |
Db 1387 AGISKELNELKGVIELLISTNYSSILEYVKKNSSESVRFSQLANGEFTKAEGEEKNASAR 1446

Qy  295 IEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILK------TNNDKTK----- 339
           :  | :||   |:  |:  |||                   :: |
Db 1447 LAEAEKLKEQIVKDL--DYSDIDDKVKKIEGIKREILKMKESALTFWEESEKFKQMCSSH 1504

Db 1505 MENAKEGKKKIEYLKNNGDGGKANITDSQME 1535
```

Search completed: March 18, 2002, 10:11:52
Job time: 970 sec

OM protein - protein search, using sw model

Copyright (c) 1993 - 2000 Compugen Ltd.

GenCore version 4.5

Run on: March 18, 2002, 10:10:54 : Search time 124.19 Seconds
(without alignments)
436.968 Million cell updates/sec

Title: US-09-596-746a-34
Perfect score: 1809
Sequence: 1 MACNNSGKDGNTSANSADES.........KNLSKAAKEMLTNSVKELTS 371

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
SPTREMBL_17:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Match Length	DB	ID	Description
1	903	49.9	194	2	O31114	borrelia bu
2	886	49.0	211	2	Q44977	borrelia bu
3	885	48.9	191	2	O31120	borrelia bu
4	884	48.9	193	2	P94242	borrelia bu
5	879	48.6	192	2	Q9S3P2	borrelia bu
6	879	48.6	210	2	O44719	borrelia bu
7	862	47.7	203	2	O08138	borrelia bu
8	849	47.1	178	2	Q44980	borrelia bu
9	790	43.7	174	2	Q9R783	borrelia bu
10	786	43.4	165	2	Q9AGB1	borrelia bu
11	782	43.2	168	2	O34121	borrelia bu
12	720.5	39.8	193	2	O31115	borrelia bu
13	664	36.7	194	2	O31122	borrelia bu
14	662	36.6	184	2	O34120	borrelia bu
15	662	36.6	193	2	P94234	borrelia bu
16	661	36.5	207	2	Q9REF7	borrelia ga
17	658.5	36.4	192	2	Q9R7B1	borrelia bu
18	658.5	36.4	193	2	Q9RRS3	borrelia bu
19	656.5	36.3	192	2	Q9S3P3	borrelia bu
20	656	36.3	200	2	O08139	borrelia bu
21	656	36.3	207	2	Q45187	borrelia bu
22	656	36.3	207	2	Q9REF8	borrelia ga
23	654	36.2	193	2	P94236	borrelia bu
24	653	36.1	207	2	Q07336	borrelia bu
25	651.5	36.0	193	2	P94231	borrelia bu
26	649.5	35.9	212	2	O44718	borrelia bu
27	649	35.9	207	2	Q9RQR9	borrelia bu
28	647.5	35.8	199	2	O08233	borrelia ta
29	647	35.8	207	2	Q45177	borrelia bu
30	646	35.7	201	2	Q9RQR8	borrelia bu
31	643	35.5	189	2	P94224	borrelia bu
32	642	35.5	191	2	P70818	borrelia bu
33	642	35.4	191	2	O44726	borrelia bu
34	638	35.3	200	2	O44996	borrelia bu
35	636.5	35.2	211	2	O44976	borrelia bu
36	635	35.1	207	2	Q49581	borrelia ga
37	633	35.1	202	2	Q9S3P0	borrelia bu
38	630	34.8	209	2	Q44883	borrelia bu
39	628.5	34.7	201	2	P96573	borrelia bu
40	628	34.7	191	2	P96572	borrelia ja
41	624.5	34.5	191	2	P96516	borrelia ja
42	623.5	34.5	201	2	O44720	borrelia ja
43	622	34.4	211	2	P94237	borrelia bu
44	621	34.3	193	2	P94237	borrelia ja
45	616.5	34.1	201	2	P96571	borrelia ja

ALIGNMENTS

RESULT 1
O31114

ID   O31114   PRELIMINARY;   PRT;   194 AA.
AC   O31114;
DT   01-JAN-1998 (TREMBLrel. 05, Created)
DT   01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT   01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C (FRAGMENT)
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=OC2;
RA   Wang I.-N., Dykhuizen D.E., Dunn J.J., Luft B.J.;
RL   Submitted (OCT-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF029861; AAB86544.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
FT   NON_TER   1   1
FT   NON_TER   194   194
SQ   SEQUENCE   194 AA; 20452 MW; 712A6C5B7873C21F CRC64;

Query Match   49.9%; Score 903; DB 2; Pred. No. 2.8e-33; Length 194;
Best Local Similarity   98.9%;
Matches 186; Conservative   2; Mismatches   0; Indels   0; Gaps   0;

Qy   1 MACNNSGKDGNTSANSADESVKGPNLTEISRKITDSNAVLLAVKEVEALLSSIDELAKAI 60
             ::||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7 ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDELAKAI 66

Qy   61 GKKIKNDGSLDNEANRNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEEFST 120
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   67 GKKIKNDGSLDNEANRNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEEFST 126

Qy   121 KLKDNHAQLGIGQVTDENAKKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANSV 180
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   127 KLKDNHAQLGIGQVTDENAKKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANSV 186

```
Qy    181 KELTSPVV 188
          ||||||||
Db    187 KELTSPVV 194

RESULT 2
ID   Q44977      PRELIMINARY;      PRT;   211 AA.
AC   Q44977;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C.
OS   OSPC.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PBRE;
RX   MEDLINE=95995018; PubMed=7655660;
RA   Jauris-Heipke S., Liegl G., Preac-Mursic V., Roessler D., Schwab E.,
RA   Soutschek E., Will G., Wilske B.;
RT   "Molecular analysis of genes encoding outer surface protein C (OspC)
RT   of Borrelia burgdorferi sensu lato: relationship to ospA genotype and
RT   evidence of lateral gene exchange of ospC.";
RL   J. Clin. Microbiol. 33:1860-1866(1995).
DR   EMBL; X81522; CAA57242.1; -.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   ProDom; PD001149; Lipoprotein_6; 1.
SQ   SEQUENCE   211 AA;  22214 MW;  BFE8C7820CDE10D3 CRC64;

Query Match             49.0%;  Score 886;  DB 2;  Length 211;
Best Local Similarity   97.3%;
Matches  183;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDELAKAI 60
          :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     17 ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDELAKAI 76

Qy     61 GKKIKNDGSLDNEANRNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEEFST 120
          |||||||||| ||| |||||||||||||||||||||||||||||||||||||||||||||
Db     77 GKKIKNDGSLGDEANHNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEEFST 136

Qy    121 KLKDNHAQLGIQGVTDENAKRAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANSV 180
          |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db    137 KLKDNHAQLGIQGVTDENAKKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANSV 196

Qy    181 KELTSPVV 188
          ||||||||
Db    197 KELTSPVV 204
```

```
RESULT 4
ID   P94242      PRELIMINARY;      PRT;   193 AA.
AC   P94242;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C (FRAGMENT).
OS   OSPC.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ZS7;
RX   MEDLINE=96296448; PubMed=8709845;
RA   Livey I., Gibbs C.P., Schuster R., Dorner F.;
RT   "Evidence for lateral transfer and recombination in OspC variation in
RT   Lyme disease Borrelia.";
RL   Mol. Microbiol. 18:257-269(1995).
DR   EMBL; L42868; AAB37011.1; -.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   ProDom; PD001149; Lipoprotein_6; 1.
SQ   SEQUENCE   193 AA;  20189 MW;  64E30C240EEF54CA CRC64;

Query Match             48.9%;  Score 885;  DB 2;  Length 191;
Best Local Similarity  100.0%;  Pred. No. 1.7e-32;
Matches  183;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    191 NNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVETLLTSIDELAKAIGKK 250
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     10 NNSGKDGNTSANSADESVKGPNLTEISKKITESNAVLLAVKEVETLLTSIDELAKAIGKK 69

Qy    251 IKNDVSLDNEADHNGSLISGAYLISNLITKKISAIRDSGELKAEIEKAKKCSEEFTAKLK 310
          |||||||||||| ||||||| |||||||||||||| |||||||||| |||||||||||||
Db     70 IKNDVSLDNEADINGSLISGAVLISNLITKKISAIKDSGELKAEIERAKKCSEEFTAKLK 129

Qy    311 GEHTDLGKEGYVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLSKAAKEMLTNSVKELT 370
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    130 GEHTDLGKEGYVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLSKAAKEMLTNSVKELT 189

Qy    371 S 371
          |
Db    190 S 190
```

```
RESULT 3
ID   O31120      PRELIMINARY;      PRT;   191 AA.
AC   O31120;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C (FRAGMENT).
OS   OSPC.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=OC10;
RA   Wang I.-N., Dykhuizen D.E., Dunn J.J., Luft B.J.;
RL   Submitted (OCT-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF029869; AAB86552.1; -.

Query Match             48.9%;  Score 884;  DB 2;  Length 193;
Best Local Similarity   98.4%;  Pred. No. 1.9e-32;
Matches  183;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      3 CNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDELAKAIGK 62
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 CNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDELAKAIGK 60

Qy     63 KIKNDGSLDNEANRNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEEFSTKL 122
          |||||||| ||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 KIKNDGSLGDEANHNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEEFSTKL 120

Qy    123 KDNHAQLGIQGVTDENAKKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANSVKE 182
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 KDNHAQLGIQGVTDENAKKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANSVKE 180
```

```
Qy    183 LTSPVV 188
          ||||||
Db    181 LTSPVV 186


RESULT 5
ID   Q9S3P2       PRELIMINARY;    PRT;   192 AA.
AC   Q9S3P2;
DT   01-MAY-2000 (TREMBLrel. 13, Created)
DT   01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C (FRAGMENT).
GN   OSPC.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=297;
RX   MEDLINE=96296448; PubMed=8709845;
RA   Livey I., Gibbs C.P., Schuster R., Dorner F.;
RT   "Evidence for lateral transfer and recombination in OspC variation in
RT   Lyme disease Borrelia.";
RL   Mol. Microbiol. 18:257-269(1995).
DR   EMBL; L42893; AAB37001.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
FT   NON_TER      1      1
FT   NON_TER    192    192
SQ   SEQUENCE   192 AA;   20472 MW;  46AC8F93E4DFED6C CRC64;

Query Match             48.6%;  Score 879;  DB 2;  Length 192;
Best Local Similarity   99.4%;
Matches  180;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;   Pred. No. 3.1e-32;

Qy    191 NNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVEVETLLTSIDELAKAIGKK 250
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2   NNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVEVETLLTSIDELAKAIGKK 61

Qy    251 IKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAEIEKAKKCSEEFTAKLK 310
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    62  IKNDVSLDNEADHNGSLISGAYLISTLITKKISAIKDSGELKAEIEKAKKCSEEFTAKLK 121

Qy    311 GEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLSKAAKEMLTNSVKELT 370
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    122 GEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLSKAAKEMLTNSVKELT 181

Qy    371 S 371
          |
Db    182 S 182


RESULT 6
ID   Q44719       PRELIMINARY;    PRT;   210 AA.
AC   Q44719;
DT   01-NOV-1996 (TREMBLrel. 01, Created)
DT   01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT   01-JUN-2001 (TREMBLrel. 17, Last annotation update)
OSPC.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=HB19;
RX   MEDLINE=95154673; PubMed=7851744;
RA   Stevenson B., Barthold S.W.;
RT   "Expression and sequence of outer surface protein C among North
RT   American isolates of Borrelia burgdorferi.";
RL   FEMS Microbiol. Lett. 124:367-372(1994).
DR   EMBL; U04281; AAC43297.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
SQ   SEQUENCE   210 AA;   2D672991D584E4EE CRC64;

Query Match             48.6%;  Score 879;  DB 2;  Length 210;
Best Local Similarity   99.4%;
Matches  180;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;   Pred. No. 3.4e-32;

Qy    191 NNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVEVETLLTSIDELAKAIGKK 250
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    20  NNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVEVETLLTSIDELAKAIGKK 79

Qy    251 IKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAEIEKAKKCSEEFTAKLK 310
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    80  IKNDVSLDNEADHNGSLISGAYLISTLITKKISAIKDSGELKAEIEKAKKCSEEFTAKLK 139

Qy    311 GEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLSKAAKEMLTNSVKELT 370
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    140 GEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLSKAAKEMLTNSVKELT 199

Qy    371 S 371
          |
Db    200 S 200


RESULT 7
ID   Q08138       PRELIMINARY;    PRT;   203 AA.
AC   Q08138;
DT   01-NOV-1996 (TREMBLrel. 01, Created)
DT   01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT   01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C PRECURSOR (FRAGMENT).
OSPC.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DK7;
RX   MEDLINE=94075528; PubMed=8253951;
RA   Theisen M., Frederiksen B., Lebech A.M., Vuust J., Hansen K.;
RT   "Polymorphism in ospC gene of Borrelia burgdorferi and
RT   immunoreactivity of OspC protein: implications for taxonomy and for
RT   use of OspC protein as a diagnostic antigen.";
RL   J. Clin. Microbiol. 31:2570-2576(1993).
CC   -!- SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID
CC       ANCHOR.
DR   EMBL; X73625; CAA52004.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
KW   Signal; Lipoprotein; Outer membrane.
FT   SIGNAL     1     18     POTENTIAL.
FT   CHAIN     19    203     OUTER SURFACE PROTEIN C.
FT   NON_TER  203    203
SQ   SEQUENCE   203 AA;   21364 MW;  AC9B29FE0E69B48F CRC64;

Query Match             47.7%;  Score 862;  DB 2;  Length 203;
Best Local Similarity   95.7%;
Matches  179;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;   Pred. No. 1.9e-31;

Qy    1  MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDELAKAI 60
            ||||||||||||||||||||||||||||||||| ||||:||||||:|:|||||||||||
Db    17 ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDELAKAI 76
```

```
Qy    61 GKKIKNDGSLDNEANRNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEEFST 120
         ||||||||| :|| ||||||||||||||||||||||||||||||||||||||||||||||
Db    77 GKKIKNDGSLGDEANHNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEEFST 136

Qy   121 KLKDNHAQLGIQGVTDENAKKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANSV 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   137 KLKDNHAQLGIQGVTDENAKKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANSV 196

Qy   181 KELTSPV 187
         |||||||
Db   197 KSLQSVV 203
```

```
RESULT 8

ID    Q44980             PRELIMINARY;    PRT;    178 AA.
AC    Q44980;
DT    01-NOV-1996 (TrEMBLrel. 01, Created)
DT    01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT    01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE    OUTER SURFACE PROTEIN C (FRAGMENT).
GN    OSPC.
OS    Borrelia burgdorferi (Lyme disease spirochete).
OC    Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX    NCBI_TaxID=139;
RN    [1]
RP    SEQUENCE FROM N.A.
RC    STRAIN=BUR;
RX    MEDLINE=95286481; PubMed=7768799;
RA    Theisen M., Borre M., Mathiesen M.J., Mikkelsen B., Lebech A.M.,
RA    Hansen K.;
RT    "Evolution of the Borrelia burgdorferi outer surface protein OspC.";
RL    J. Bacteriol. 177:3036-3044(1995).
DR    EMBL; X84765; CAA59236.1; -.
DR    InterPro; IPR001800; Lipoprotein_6.
DR    Pfam; PF01441; Lipoprotein_6; 1.
DR    ProDom; PD001149; Lipoprotein_6; 1.
FT    NON_TER     1      1
FT    NON_TER   178    178
SQ    SEQUENCE  178 AA;  18724 MW;  49A6F785145CC208 CRC64;

Query Match             46.9%;  Score 849;  DB 2;  Length 178;
Best Local Similarity   99.4%;  Pred. No. 6e-31;
Matches  177;  Conservative   1;  Mismatches   0;  Indels   0;  Gaps   0;

Qy    11 NTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDELAKAIGKKIKNDGSL 70
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 NTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDELAKAI 49

Qy    71 DNEANRNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEEFSTKLKDNHAQLG 130
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 DNEANRNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEEFSTKLKDNHAQLG 120

Qy   131 IQGVTDENAKKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANSVKELTSPVV 188
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 IQGVTDENAKKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANSVKELTSPVV 178
```

```
RESULT 9

ID    Q9R7B3             PRELIMINARY;    PRT;    174 AA.
AC    Q9R7B3;
DT    01-MAY-2000 (TrEMBLrel. 13, Created)
DT    01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT    01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE    OUTER SURFACE PROTEIN C (FRAGMENT).
OS    Borrelia burgdorferi (Lyme disease spirochete).
OC    Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX    NCBI_TaxID=139;
RN    [1]
RP    SEQUENCE FROM N.A.
RC    STRAIN=TETS;
RX    MEDLINE=97478003; PubMed=9336916;
RA    Ras N.M., Postic D., Foretz M., Baranton G.;
RT    "Borrelia burgdorferi sensu stricto, a bacterial species 'made in the
RL    U.S.A. '?";
RL    Int. J. Syst. Bacteriol. 47:1112-1117(1997).
RN    [2]
RP    SEQUENCE FROM N.A.
RC    STRAIN=TETS;
RA    Marti-Ras N., Postic D., Foretz M., Baranton G.;
RL    Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.
DR    EMBL; U91795; AAB81892.1; -.
DR    InterPro; IPR001800; Lipoprotein_6.
DR    Pfam; PF01441; Lipoprotein_6; 1.
DR    ProDom; PD001149; Lipoprotein_6; 1.
SQ    SEQUENCE  174 AA;  18291 MW;  A14F5B8BA151A278 CRC64;

Query Match             43.7%;  Score 790;  DB 2;  Length 174;
Best Local Similarity   97.0%;  Pred. No. 2.3e-28;
Matches  162;  Conservative   3;  Mismatches   2;  Indels   0;  Gaps   0;

Qy     1 MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDELAKAI 60
         ::||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     8 ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDELAKAI 67

Qy    61 GKKIKNDGSLDNEANRNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEEFST 120
         ||||||||| :|| ||||||||||||||||||||||||||||||||||||||||||||||
Db    68 GKKIKNDGSLGDEANHNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEEFST 127

Qy   121 KLKDNHAQLGIQGVTDENAKKAILKANAAGKDKGVEELEKLSGSLES 167
         |||||||||||||||||||||||||||||||||||||||||||||||
Db   128 KLKDNHAQLGIQGVTDENAKKAILKANAAGKDKGVEELEKLSGSLES 174
```

```
RESULT 10

ID    Q9AGB1             PRELIMINARY;    PRT;    165 AA.
AC    Q9AGB1;
DT    01-JUN-2001 (TrEMBLrel. 17, Created)
DT    01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT    01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE    OSPC-MODIFIED.
OS    Borrelia burgdorferi (Lyme disease spirochete).
OC    Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX    NCBI_TaxID=139;
RN    [1]
RP    SEQUENCE FROM N.A.
RC    STRAIN=HB19;
RA    Kumaran D., Eswaramoorthy S., Luft B.J., Koide S., Dunn J.J.,
RA    Lawson C.L., Swaminathan S.;
RT    "Crystal structure of outer surface protein C (OspC) from the Lyme
RT    disease spirochete, Borrelia burgdorferi.";
RL    EMBO J. 0:0-0(2001).
DR    EMBL; AF337548; AAK21289.1; -.
FT    VARIANT    61     61       M -> I.
SQ    SEQUENCE  165 AA;  17833 MW;  56E0153602F61BF CRC64;

Query Match             43.4%;  Score 786;  DB 2;  Length 165;
Best Local Similarity   98.2%;  Pred. No. 3.2e-28;
Matches  161;  Conservative   2;  Mismatches   1;  Indels   0;  Gaps   0;

Qy   208 VKGPNLTEISKKITESNAVVLAVKEVETLLTSIDELAKARIGKIKNDVSLDNEADHNGSL 267
         ::|||||||||||||:||||||||||||||||||||||||||||||||||||||||||||
Db     1 MKGPNLTEISKKITESNAVVLAVKEVETLLTSIDELAKARIGKIKNDVSLDNEADHNGSL 60

Qy   268 ISGAYLISNLIITKKISAIKDSGELKAEIEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAK 327
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 MSGAYLISTLIITKKISAIKDSGELKAEIEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAK 120
```

```
Qy   328 KAILKTNNDKTKGADELEKLFESVKNLSKAAKEMLTNSVKELTS 371
         |||||||||||||||||||||||||||||||||||||||||||
Db   121 KAILKTNNDKTKGADELEKLFESVKNLSKAAKEMLTNSVKELTS 164
```

RESULT 11
```
ID   O34121
AC   O34121;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C (FRAGMENT).
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TETS;
RX   MEDLINE=97478003; PubMed=9336916;
RA   Ras N.M., Postic D., Foretz M., Baranton G.;
RT   "Borrelia burgdorferi sensu stricto, a bacterial species 'made in the
RT   U.S.A.'?";
RL   Int. J. Syst. Bacteriol. 47:1112-1117(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TETS;
RA   Marti-Ras N., Postic D., Foretz M., Baranton G.;
RL   Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U91794; AAB18891.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
FT   NON_TER       1
FT   NON_TER     168
SQ   SEQUENCE   168 AA;  17704 MW;  CDA45DD8A49C4409 CRC64;
                        PRELIMINARY;    PRT;    168 AA.
```

```
Query Match             43.2%;  Score 782;  DB 2;  Length 168;
Best Local Similarity   98.2%;
Matches  164;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;
                                   Pred. No. 5e-28;
```

```
Qy    19 ESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDELAKAIGKKIKNDGSLDNEANRNE 78
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 ESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDELAKAIGKKIKNDGSLGDEANHNE 60

Qy    79 SLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEEFSTKLKDNHAQLGIQGVTDEN 138
         ||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db    61 SLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEEFSTKLKDNHAQLGIQGVTDEN 120

Qy   139 AKKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANSVKELTS 185
         |||||||||||||||||||||||||||||||||||||||||||||
Db   121 ARKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANSVKELTS 167
```

RESULT 12
```
ID   O31115
AC   O31115;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C (FRAGMENT).
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=OC3;
RA   Wang I.-N., Dykhuizen D.E., Dunn J.J., Luft B.J.;
RL   Submitted (OCT-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF029862; AAB86545.1; -.
```

```
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
FT   NON_TER       1
FT   NON_TER     193
SQ   SEQUENCE   193 AA;  20596 MW;  018A4CB310475A58 CRC64;
                        PRELIMINARY;    PRT;    193 AA.
```

RESULT 13
```
ID   O31122
AC   O31122;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C (FRAGMENT).
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=OC12;
RA   Wang I.-N., Dykhuizen D.E., Dunn J.J., Luft B.J.;
RL   Submitted (OCT-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF029871; AAB86554.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
FT   NON_TER       1
FT   NON_TER     194
SQ   SEQUENCE   194 AA;  20640 MW;  00A5E6E2D2CE0F7F CRC64;
                        PRELIMINARY;    PRT;    194 AA.
```

```
Query Match             39.8%;  Score 720.5;  DB 2;  Length 193;
Best Local Similarity   79.8%;                Pred. No. 2.9e-25;
Matches  150;  Conservative   16;  Mismatches   21;  Indels    1;  Gaps    1;
```

```
Qy     1 MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDELAKAI 60
         ::||||||||| :||||||||||||||||||||:||||||:|||||||:||||||||||
Db     7 ISCNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELAKAI 66

Qy    61 GKKIKNDGSLDNEANRNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEEFST 120
         ||||||| ||||||| :||| :|||||||||||||||||||||||||:|| |||||||:
Db    67 GKKIKNDVSLDNEADNNGSLIAGAYTISTLITQKLSKLNGSEGLKEKIAEAKKCSEEFTK 126

Qy   121 KLKDNHAQLGIQGVTDENAKKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANSV 180
         |||::| : |:||  : || ||:||||  |:||||| ||||||:||||:|||||||:||||
Db   127 KLKEKHTDLGKKDAQTDVHAKEAILKTNGT-KDKGAAELEKLFESVENLAKAAKEMLSNSV 185

Qy   181 KELTSPVV 188
         |||||:|
Db   186 KELTSPVV 193
```

```
Query Match             36.7%;  Score 664;  DB 2;  Length 194;
Best Local Similarity   75.1%;              Pred. No. 9.1e-23;
Matches  142;  Conservative   19;  Mismatches   26;  Indels    2;  Gaps    2;
```

```
Qy     1 MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDELA-KA 59
         ::|||||||||||||||||||||||||||||||||::||||:||||:|:|||||||| ||
Db     7 ISCNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDELATKA 66

Qy    60 IGKKIKNDGSLDNEANRNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEEFS 119
         |||||:  | ||::|:| ||:|||||||| ||||||   |:::|::|:||| ||||||:|
Db    67 IGKKIQQNGGLAVEAGHNGTLIAGAYTISKLITQKLDLGLKNSEKLIENAKKCSEDFT 126

Qy   120 TKLKDNHAQLGIQGVTDENAKKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANS 179
         ||||:| |:|||||:|||||||||| :||  |:||| ||||||:| ||::|||||||||||
Db   127 KKLEGEHAQLGIENVTDENAKKAILTFDAA-KDKGAAELEKLFKAVENLAKAAKEMLANS 185

Qy   180 VKELTSPVV 188
         ||||||:|
Db   188 VKELTSPVV ...
```

```
Db      186 VKELTSPIV 194

RESULT 14
ID   O34120       PRELIMINARY;      PRT;   184 AA.
AC   O34120;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C (FRAGMENT).
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97478003; PubMed=9336916;
RA   Ras N.M., Postic D., Foretz M., Baranton G.;
RT   "Borrelia burgdorferi sensu stricto, a bacterial species 'made in the
RT   U.S.A.'";
RL   Int. J. Syst. Bacteriol. 47:1112-1117(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TETS;
RA   Marti-Ras N., Postic D., Foretz M., Baranton G.;
RL   Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U91793; AAB81890.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
FT   NON_TER       1      1
FT   NON_TER     184    184
SQ   SEQUENCE   184 AA;   19391 MW;  9B1C12B79273FB0E CRC64;

Query Match            36.6%;  Score 662;  DB 2;  Length 184;
Best Local Similarity  76.3%;  Pred. No. 1e-22;
Matches  142;  Conservative  15;  Mismatches  27;  Indels  2;  Gaps  2;

Qy    3 CNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDELAKAIGK 62
         ||||||| ||||||||||||||||||||||||||||:||||||||||:|||||||||||
Db    1 CNNSGKDGNASANSADESVKGPNLTEISKKITDSNAFVLAVKEVEALISSIDELAKAIGK 60

Qy   63 KIKNDGSLDNEANRNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEEFSTKL 122
         :|::|||||||||||||||||:||:|||||||||:|:|||||||:||||||||| |:||
Db   61 KVEANG-LGNEADRNTSLLAGAHEISILITQKLTALKDSGGLKAETAEAKKCSEAFTKKL 119

Qy  123 KDNHAQLGIQGVTDENAKKAILKANAAGKDKGVEELEKLSGGSLESLSKAAKEMLANSVKE 182
         ||||| |||||||| ||||||| |  |||||||||||||||||||||||||||||||||
Db  120 KDNHAQFGIQNVQDVEAKKAILKTNAT-KDKGAEDLEKLFKAVESLSKAAQEALTNSVKE 178

Qy  183 LTSPVV 188
         ||:|||
Db  179 LTNPVV 184
```

```
RA   Livey I., Gibbs C.P., Schuster R., Dorner F.;
RT   "Evidence for lateral transfer and recombination in OspC variation in
RT   Lyme disease Borrelia.";
RL   Mol. Microbiol. 18:257-269(1995).
DR   EMBL; L42895; AAB37003.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
FT   NON_TER       1      1
FT   NON_TER     193    193
SQ   SEQUENCE   193 AA;   20411 MW;  05B68720F061E2A0 CRC64;

Query Match            36.6%;  Score 662;  DB 2;  Length 193;
Best Local Similarity  75.9%;  Pred. No. 1.1e-22;
Matches  142;  Conservative  17;  Mismatches  26;  Indels  2;  Gaps  2;

Qy    3 CNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDELA-KAIG 61
         ||||||||||||||||||||||||||||||||:||||:||||||:|:||:|||||| |||
Db    1 CNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDELATKAIG 60

Qy   62 KKIKNDGSLDNEANRNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEEFSTK 121
         ||:|:|::  |:||  ||||||||||:||||||| |:||:|||||:|| |||||:|:|
Db   61 KKIQQNGLAVEAGHNGTLLAGAYTISKLIITQKLDGLKNSEKLKEKIENAKKCSEDFTKK 120

Qy  122 LKDNHAQLGIQGVTDENAKKAILKANAAGKDKGVEELEKLSGGSLESLSKAAKEMLANSVK 181
         | ||||||||:|||||||||:||| |  |||||| |:|||||:|::|:|||||||||||||
Db  121 LEGEHAQLGIENVTDENAKKALLITDAA-KDKGAAELEKLFKAVENLAKAARKEMLANSVK 179

Qy  182 ELTSPVV 188
         ||||||:|
Db  180 ELTSPIV 186
```

Search completed: March 18, 2002, 10:10:54
Job time: 977 sec

```
RESULT 15
ID   P94234       PRELIMINARY;      PRT;   193 AA.
AC   P94234;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C (FRAGMENT).
OSPC.
GN   OSPC.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96296448; PubMed=8709845;
```

THIS PAGE BLANK (USPTO)

OM protein - protein search, using sw model

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

Run on: March 18, 2002, 09:54:32 ; Search time 118.14 Seconds

Title: US-09-596-746A-38
Perfect score: 1797
Sequence: 1 MACNNSGKDGNASANSADES.........ENLSKAAQDTLKNAVKELTS 368

Scoring table: BLOSUM62
             Gapop 10.0 , Gapext 0.5
Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                Maximum Match 100%
                Listing first 45 summaries

Database :   A_Geneseq_1101:*

1:	/SIDS8/gcgdata/geneseq/geneseqp/AA1980.DAT:*
2:	/SIDS8/gcgdata/geneseq/geneseqp/AA1981.DAT:*
3:	/SIDS8/gcgdata/geneseq/geneseqp/AA1982.DAT:*
4:	/SIDS8/gcgdata/geneseq/geneseqp/AA1983.DAT:*
5:	/SIDS8/gcgdata/geneseq/geneseqp/AA1984.DAT:*
6:	/SIDS8/gcgdata/geneseq/geneseqp/AA1985.DAT:*
7:	/SIDS8/gcgdata/geneseq/geneseqp/AA1986.DAT:*
8:	/SIDS8/gcgdata/geneseq/geneseqp/AA1987.DAT:*
9:	/SIDS8/gcgdata/geneseq/geneseqp/AA1988.DAT:*
10:	/SIDS8/gcgdata/geneseq/geneseqp/AA1989.DAT:*
11:	/SIDS8/gcgdata/geneseq/geneseqp/AA1990.DAT:*
12:	/SIDS8/gcgdata/geneseq/geneseqp/AA1991.DAT:*
13:	/SIDS8/gcgdata/geneseq/geneseqp/AA1992.DAT:*
14:	/SIDS8/gcgdata/geneseq/geneseqp/AA1993.DAT:*
15:	/SIDS8/gcgdata/geneseq/geneseqp/AA1994.DAT:*
16:	/SIDS8/gcgdata/geneseq/geneseqp/AA1995.DAT:*
17:	/SIDS8/gcgdata/geneseq/geneseqp/AA1996.DAT:*
18:	/SIDS8/gcgdata/geneseq/geneseqp/AA1997.DAT:*
19:	/SIDS8/gcgdata/geneseq/geneseqp/AA1998.DAT:*
20:	/SIDS8/gcgdata/geneseq/geneseqp/AA1999.DAT:*
21:	/SIDS8/gcgdata/geneseq/geneseqp/AA2000.DAT:*
22:	/SIDS8/gcgdata/geneseq/geneseqp/AA2001.DAT:*

(without alignments)
230.734 Million cell updates/sec

SUMMARIES

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1797	100.0	368	22	AAB62717	Borrelia sp chimer
2	1789	99.6	391	22	AAB62736	Borrelia sp chimer
3	1533	85.3	367	22	AAB62719	Borrelia sp chimer
4	1530	85.1	391	22	AAB62735	Borrelia sp chimer
5	1485.5	82.7	368	22	AAB62718	Borrelia sp chimer
6	1482.5	82.5	392	22	AAB62734	Borrelia sp chimer
7	1479	82.3	369	22	AAB62714	Borrelia sp chimer
8	1476	82.1	393	22	AAB62732	Borrelia sp chimer
9	1225.5	68.2	377	22	AAB62713	Borrelia sp chimer
10	1224	68.1	373	22	AAB62711	Borrelia sp chimer
11	1221	67.9	397	22	AAB62729	Borrelia sp chimer

ALIGNMENTS

RESULT 1
AAB62717

ID   AAB62717 standard; protein; 368 AA.
XX
AC   AAB62717;
XX
DT   03-APR-2001 (first entry)
XX
DE   Borrelia sp chimeric ospC protein SEQ ID NO: 38.
KW   Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
OS   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
XX
PN   WO200078966-A1.
XX
PD   28-DEC-2000.
XX
PF   19-JUN-2000; 2000WO-US16915.
XX
PR   18-JUN-1999; 99US-0140042.
XX
PA   (UYNY) UNIV NEW YORK STATE RES FOUND.
PA   (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI   Dattwyler RJ, Seinost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
DR   WPI; 2001-050113/06.
DR   N-PSDB; AAF29021.
XX
PT   Compositions of OspC polypeptides from strains of Borrelia which cause
PT   Lyme disease are used to immunize animals and detect immune responses
PT   to Lyme disease -

12	1220	67.9	369	22	AAB62716	Borrelia sp chimer
13	1219.5	67.9	400	22	AAB62739	Borrelia sp chimer
14	1217	67.7	393	22	AAB62731	Borrelia sp chimer
15	1178	65.6	378	22	AAB62712	Borrelia sp chimer
16	1176.5	65.5	374	22	AAB62710	Borrelia sp chimer
17	1173.5	65.3	398	22	AAB62728	Borrelia sp chimer
18	1172.5	65.2	370	22	AAB62715	Borrelia sp chimer
19	1172	65.2	401	22	AAB62738	Borrelia sp chimer
20	1169.5	65.1	394	22	AAB62730	Borrelia sp chimer
21	1167.5	65.0	560	22	AAB62724	Borrelia sp chimer
22	1157.5	64.4	384	22	AAB62726	Borrelia sp chimer
23	1154.5	64.2	408	22	AAB62737	Borrelia sp chimer
24	1142	63.6	378	22	AAB62725	Borrelia sp chimer
25	1134	63.1	401	22	AAB62733	Borrelia sp chimer
26	1099.5	61.2	410	22	AAB62740	Borrelia sp chimer
27	1098.5	61.1	386	22	AAB62727	Borrelia sp chimer
28	901	50.1	191	22	AAB62705	B burgdorferi ospC
29	899	50.0	184	22	AAB62702	B burgdorferi ospC
30	899	50.0	191	22	AAR60884	Borrelia 28691 O
31	889	49.5	191	22	AAB62706	B burgdorferi ospC
32	673	37.5	466	16	AAR75740	B31 outer surface
33	663	36.9	193	22	AAB62709	B burgdorferi ospC
34	657.5	36.6	212	16	AAW11934	Outer surface prot
35	651.5	36.3	194	15	AAR75739	B31 outer surface
36	640	35.6	587	16	AAR60885	B31 Osp-A/antigen
37	638.5	35.5	466	16	AAB62709	B burgdorferi ospC
38	623.5	34.7	192	22	AAB62703	B burgdorferi ospC
39	623.5	34.7	210	16	AAR75746	B burgdorferi ospC
40	622.5	34.6	192	22	AAB62721	Outer surface prot
41	621.5	34.5	192	15	AAR60886	B burgdorferi IP2 OspC
42	620.5	34.5	209	22	AAB62720	B burgdorferi ospC
43	617.5	34.4	210	16	AAR75727	B. burgdorferi str
44	608.5	33.9	178	15	AAR62771	B. burgdorferi 2591 anti
45	605.5	33.7	194	15	AAR60894	Borrelia H9 antige

```
XX
PS   Claim 43; Page 97; 160pp; English.
XX
CC   The present invention provides compositions comprising ospC proteins and
CC   chimeric ospC proteins from members of the Borrelia genus. These may be
CC   Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC   vaccines against Borrelia infection, which is spread by ticks and leads
CC   to Lyme disease.
XX
SQ   Sequence   368 AA;

Query Match             100.0%;  Score 1797;  DB 22;  Length 368;
Best Local Similarity   100.0%;  Pred. No. 1.5e-107;
Matches  368;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    1 MACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKA 60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 macnnsgkdgnasansadesvkgpnlteiskkitesnavvlavkevetllasidelatka 60

Qy   61 IGKKIIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNK 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 igkkiignngleanqskntsllsgayaisdliaeklnvlkneelkekidtakqcsteftnk 120

Qy  121 LKSEHAVLGLDNLTDDNAQRAILKKHANKDGAAELEKLFKAVENLSKAAQDTLKNAVKE 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 lksehavlgldnltddnaqrailkkhankdgaaelekifkavenlskaaqdtlknavke 180

Qy  181 LTSPIVHGNNSRKDGNASTNSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDE 240
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 ltspivhgnnsrkdgnastnsadesvkgpnlteiskkitesnavvlavkevetllaside 240

Qy  241 LATKAIGKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCST 300
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 latkaigkkignngleanqskntsllsgayaisdliaeklnvlkneelkekidtakqcst 300

Qy  301 EFTNKLKSEHAVLGLDNLTDDNAQRAILKKHANKDGAAELEKLFKAVENLSKAAQDTLK 360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 eftnklksehavlgldnltddnaqrailkkhankdgaaelekifkavenlskaaqdtlk 360

Qy  361 NAVKELTS 368
        ||||||||
Db  361 navkelts 368


RESULT   2
AAB62736
ID   AAB62736 standard; Protein; 391 AA.
XX
AC   AAB62736;
XX
DT   03-APR-2001 (first entry)
XX
DE   Borrelia sp chimeric ospc protein SEQ ID NO: 76.
XX
KW   Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
XX
PN   WO200078966-A1.
XX
PD   28-DEC-2000.
XX
PF   19-JUN-2000; 2000WO-US16915.
XX
PR   18-JUN-1999; 99US-0140042.
XX
PA   (UYNY ) UNIV NEW YORK STATE FES FOUND.
PA   (BROO-) BROOK BIOTECHNOLOGIES INC.
PI   Dattwyler RJ, Seinost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
```

```
DR   WPI; 2001-050113/06.
DR   N-PSDB; AAF29040.
XX
PT   Lyme disease which cause
PT   to Lyme disease -
XX
CC   The present invention provides compositions comprising ospC proteins and
CC   chimeric ospC proteins from members of the Borrelia genus. These may be
CC   Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC   vaccines against Borrelia infection, which is spread by ticks and leads
CC   to Lyme disease.
XX
PS   Claim 43; Page 144-145; 160pp; English.
XX
SQ   Sequence   391 AA;

Query Match             99.6%;  Score 1789;  DB 22;  Length 391;
Best Local Similarity   99.7%;  Pred. No. 5.1e-107;
Matches  366;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    2 ACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKAI 61
        :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   25 acnnsgkdgnasansadesvkgpnlteiskkitesnavvlavkevetllasidelatkai 84

Qy   62 GKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNKL 121
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   85 gkkignngleanqskntsllsgayaisdliaeklnvlkneelkekidtakqcsteftnkl 144

Qy  122 KSEHAVLGLDNLTDDNAQRAILKKHANKDGAAELEKLFKAVENLSKAAQDTLKNAVKEL 181
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  145 ksehavlgldnltddnaqrailkkhankdgaaelekifkavenlskaaqdtlknavkel 204

Qy  182 TSPIVHGNNSRKDGNASTNSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDEL 241
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  205 tspivhgnnsrkdgnastnsadesvkgpnlteiskkitesnavvlavkevetllasidel 264

Qy  242 ATKAIGKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTE 301
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  265 atkaigkkignngleanqskntsllsgayaisdliaeklnvlkneelkekidtakqcste 324

Qy  302 FTNKLKSEHAVLGLDNLTDDNAQRAILKKHANKDGAAELEKLFKAVENLSKAAQDTLKN 361
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  325 ftnklksehavlgldnltddnaqrailkkhankdgaaelekifkavenlskaaqdtlkn 384

Qy  362 AVKELTS 368
        |||||||
Db  385 avkelts 391


RESULT   3
AAB62719
ID   AAB62719 standard; Protein; 367 AA.
XX
AC   AAB62719;
XX
DT   03-APR-2001 (first entry)
XX
DE   Borrelia sp chimeric ospc protein SEQ ID NO: 42.
XX
KW   Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
XX
PN   WO200078966-A1.
XX
PD   28-DEC-2000.
XX
PF   19-JUN-2000; 2000WO-US16915.
XX
PR   18-JUN-1999; 99US-0140042.
XX
```

```
XX
PA   (UYNY ) UNIV NEW YORK STATE RES FOUND.
PA   (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
XX
PI   Dattwyler RJ, Selnost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
XX
DR   WPI; 2001-050113/06.
DR   N-PSDB; AAF29023.
XX
XX
PT   Compositions of OspC polypeptides from strains of Borrelia which cause
PT   Lyme disease are used to immunize animals and detect immune responses
PT   to Lyme disease -
XX
PS   Claim 43; Page 102-103; 160pp; English.
XX
CC   The present invention provides compositions comprising ospc proteins and
CC   chimeric ospC proteins from members of the Borrelia genus. These may be
CC   Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC   vaccines against Borrelia infection, which is spread by ticks and leads
CC   to Lyme disease.
XX
SQ   Sequence    367 AA;

Query Match       85.3%;  Score 1533;  DB 22;  Length 367;
Best Local Similarity    86.6%;  Pred. No. 1.1e-90;
Matches  318;  Conservative  20;  Mismatches  27;  Indels   2;  Gaps   2;

Qy  2   ACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVEVETLLASIDELATKAI  61
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1   acnnsgkdgnasansadesvkgpnlteiskkitesnavvlavkevevetllasidelatkai  60

Qy  62  GKKIGNGNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNKL  121
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  61  gkkigngngleangskntslisgayaisdliaeklnvlkneelkekidtakqcsteftnkl  120

Qy  122 KSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKEL  181
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 ksehavlgldnltddnaqrailkkhankdkgaaeleklfkavenlskaaqdtlknavkel  180

Qy  182 TSPIVHGNNSRKDGNASTNSADESVKGPNLTEISKKITESNAVVLAVKEVEVETLLASIDEL  241
        |||||||||| :|||||||||||||||||||||||||||||||||||||| :|||||||||
Db  181 tspivhgnnsgkdgntsansadesvkgpnlteiskkitesnavvlavkeveietllasidel  240

Qy  242 ATKAIGKKI-GNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKN-EELKEKIDTAKQCS  299
        |||||||||  :||||: ||||||||||||||||||||||||||| ::||:|||||||||
Db  241 atkaigkkiqgngglaveaghngtllagaytiskiitqkldglknsekienakkcs  300

Qy  300 TEFTNKLKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTL  359
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 edfkkllegehaqlgienvtdenakkallitdaakdkgaaeleklfkavenlakaakeml  360

Qy  360 KNAVKEL  366
        |:||||
Db  361 ansvkel  367

RESULT  4
AAB62735
ID   AAB62735 standard; protein; 391 AA.
XX
AC   AAB62735;
XX
DT   03-APR-2001  (first entry)
XX
DE   Borrelia sp chimeric ospc protein SEQ ID NO: 74.
XX
KW   Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
XX
PN   WO200078966-A1.
```

```
XX
PD   28-DEC-2000.
XX
PF   19-JUN-2000; 2000WO-US16915.
XX
PR   18-JUN-1999; 99US-0140042.
XX
PA   (UYNY ) UNIV NEW YORK STATE RES FOUND.
PA   (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI   Dattwyler RJ, Selnost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
DR   WPI; 2001-050113/06.
DR   N-PSDB; AAF29039.
XX
PT   Compositions of OspC polypeptides from strains of Borrelia which cause
PT   Lyme disease are used to immunize animals and detect immune responses
PT   to Lyme disease -
XX
PS   Claim 43; Page 142; 160pp; English.
XX
CC   The present invention provides compositions comprising ospc proteins and
CC   chimeric ospC proteins from members of the Borrelia genus. These may be
CC   Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC   vaccines against Borrelia infection, which is spread by ticks and leads
CC   to Lyme disease.
XX
SQ   Sequence    391 AA;

Query Match       85.1%;  Score 1530;  DB 22;  Length 391;
Best Local Similarity    86.4%;  Pred. No. 1.8e-90;
Matches  317;  Conservative  21;  Mismatches  27;  Indels   2;  Gaps   2;

Qy  2   ACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVEVETLLASIDELATKAI  61
        : ||||||||||||||||||||||||||||||||||||||||||| :|||||||||||||
Db  1   scnnsgkdgnasansadesvkgpnlteiskkitesnavvlavkeveietllasidelatkai  60

Qy  62  GKKIGNGNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNKL  121
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  25  gkkignngleangsknts...                                  84

Qy  122 KSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKEL  181
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  85  gkkignngleangsknts... gayaisdliaeklnvlkneelkekidtakqcsteftnkl  144

Qy  182 TSPIVHGNNSRKDGNASTNSADESVKGPNLTEISKKITESNAVVLAVKEVEVETLLASIDEL  241
        ||||||||||  :||||||||||||||||||||||||||||||||| :||||||||
Db  145 ksehavlgldnltddnaqrailkkhankdkgaaeleklfkavenlskaaqdtlknavkel  204

Qy  242 ATKAIGKKI-GNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKN-EELKEKIDTAKQCS  299
        |||||||||  :||||: ||||||||||||||||||||||||||| ::||:|||||||||
Db  204 atkaigkkiqgngglaveaghngtllagaytiskiitqkldglknsekienakkcs  264

Qy  300 TEFTNKLKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTL  359
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  265 atkaigkkiqgngglaveaghngtliagaytiskitqkldglknsekikenakkcs  324

Qy  300 TEFTNKLKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTL  359
        |||||||||||||||||||||||||||||||||||||||||||||| :|||||||: :
Db  265 tspivhgnnsgkdgntsansadesvkgpnlteiskkitesnavvlavkeveietllasidel  324

Qy  360 KNAVKEL  366
        |:||||
Db  325 edftkkllegehaqlgienvtdenakkalitdaakdkgaaeleklfkavenlakaakeml  384

RESULT  5
AAB62718
ID   AAB62718 standard; protein; 368 AA.
XX
AC   AAB62718;
XX
DT   03-APR-2001  (first entry)
XX
DE   Borrelia sp chimeric ospc protein SEQ ID NO: 40.
```

```
Qy  300 TEFTNKLKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKEL  384
Db  325 edftkkllegehaqlgienvtdenakkalitdaakdkgaaeleklfkavenlakaakeml
Qy  360 KNAVKEL  366
Db  385 ansvkel  391
```

Left column

```
RESULT    6
AAB62734
ID   AAB62734 standard; Protein; 392 AA.
XX
KW   Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
XX
PN   WO200078966-A1.
XX
PD   28-DEC-2000.
XX
PF   19-JUN-2000; 2000WO-US16915.
XX
PR   18-JUN-1999;    99US-0140042.
XX
PA   (BROO-) BROOK BIOTECHNOLOGIES INC.
PA   (UYNY ) UNIV NEW YORK STATE RES FOUND.
XX
PI   Dattwyler RJ,  Seinost G,  Dykhuizen D,  Luft BJ,  Gomes-Solecki M;
XX
DR   N-PSDB; AAF29022.
PR   WPI; 2001-050113/06.
XX
PT   to Lyme disease.
PT   Lyme disease are used to immunize animals and detect immune responses
PT   Compositions of Ospc polypeptides from strains of Borrelia which cause
XX
PS   Claim 43; Page 99-100; 160pp; English.
XX
CC   to Lyme disease.
CC   vaccines against Borrelia infection, which is spread by ticks and leads
CC   Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC   chimeric ospC proteins from members of the Borrelia genus. These may be
CC   The present invention provides compositions comprising ospc proteins and
XX
SQ   Sequence   368 AA;

Query Match             82.7%;  Score 1485.5;  DB 22;  Length 368;
  Best Local Similarity   84.3%;
  Matches  311;  Conservative   24;  Mismatches   31;  Indels    3;  Gaps    3;
                                        Pred. No. 1.2e-87;

Qy           2 ACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKAI 61
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 acnnsgkdgnasansadesvkgpnlteiskkitesnavvlavkevetllasidelatkai 60

Qy          62 GKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNKL 121
               |||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db          61 gkkignngleanqskntsllsgayaisdliaeklnvlkneelkekidtakqcsteftnkl 120

Qy         122 KSEHAVLGLDNLTDDNAQRAILKKHANKDGAAELEKLFKAVENLSKAAQDTLKNAVKEL 181
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 ksehavlgldnltddnaqrailkkhankdgaaelekifkavenlskaaqdtlknavkel 180

Qy         182 TSPIVHGNNSRKDGNASTNSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDEL 241
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 tspivhgnnsrkdgnastnsadesvkgpnlteiskkitesnavvlavkevetllasidel 240

Qy         242 ATKAIGKKIGNN-GLEANQSKNTSLLSGAYAISDLIAEKLNVLKNE-ELKEKIDTAKQCS 299
               ||||||||||| ||||||||||||||||||||||||||||||||| |||||||||||||
Db         241 a-kaigkkikndvsldneadhngslisgaylisnlltkkisaikdsgelkaeiekakkcs 299

Qy         300 TEFTNKLKSEHAVLGLDNLTDDNAQRAILKKHANKDGAAELEKLFKAVENLSKAAQDTL 359
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         300 eeftaklkgehtdlgkegvtddnakkailktnndktkgadelekifesvknlskaakem 359

Qy         360 KNAVKELTS 368
               ||||||||
Db         360 tnsvkelts 368
```

Right column

```
AC   AAB62734;
XX
DT   03-APR-2001  (first entry)
XX
DE   Borrelia sp chimeric ospc protein SEQ ID NO: 72.
XX
KW   Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
XX
PN   WO200078966-A1.
XX
PD   28-DEC-2000.
XX
PF   19-JUN-2000; 2000WO-US16915.
XX
PR   18-JUN-1999;    99US-0140042.
XX
PA   (UYNY ) UNIV NEW YORK STATE RES FOUND.
PA   (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI   Dattwyler RJ,  Seinost G,  Dykhuizen D,  Luft BJ,  Gomes-Solecki M;
XX
DR   N-PSDB; AAF29038.
PR   WPI; 2001-050113/06.
XX
PT   to Lyme disease.
PT   Lyme disease are used to immunize animals and detect immune responses
PT   Compositions of Ospc polypeptides from strains of Borrelia which cause
XX
PS   Claim 43; Page 139-140; 160pp; English.
XX
CC   to Lyme disease.
CC   vaccines against Borrelia infection, which is spread by ticks and leads
CC   Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC   chimeric ospC proteins from members of the Borrelia genus. These may be
CC   The present invention provides compositions comprising ospc proteins and
XX
SQ   Sequence   392 AA;

Query Match             82.5%;  Score 1482.5;  DB 22;  Length 392;
  Best Local Similarity   84.0%;
  Matches  310;  Conservative   25;  Mismatches   31;  Indels    3;  Gaps    3;
                                        Pred. No. 2e-87;

Qy           2 ACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKAI 61
               :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          25 scnnsgkdgnasansadesvkgpnlteiskkitesnavvlavkevetllasidelatkai 84

Qy          62 GKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNKL 121
               |||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db          85 gkkignngleanqskntsllsgayaisdliaeklnvlkneelkekidtakqcsteftnkl 144

Qy         122 KSEHAVLGLDNLTDDNAQRAILKKHANKDGAAELEKLFKAVENLSKAAQDTLKNAVKEL 181
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         145 ksehavlgldnltddnaqrailkkhankdgaaelekifkavenlskaaqdtlknavkel 204

Qy         182 TSPIVHGNNSRKDGNASTNSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDEL 241
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         205 tspivhgnnsrkdgnastnsadesvkgpnlteiskkitesnavvlavkevetllasidel 264

Qy         242 ATKAIGKKIGNN-GLEANQSKNTSLLSGAYAISDLIAEKLNVLKNE-ELKEKIDTAKQCS 299
               ||||||||||| ||||||||||||||||||||||||||||||||| ||:|||||:||
Db         265 a-kaigkkikndvsldneadhngslisgaylisnlltkkisaikdsgelkaeiekakkcs 323

Qy         300 TEFTNKLKSEHAVLGLDNLTDDNAQRAILKKHANKDGAAELEKLFKAVENLSKAAQDTL 359
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||:|
Db         324 eeftaklkgehtdlgkegvtddnakkailktnndktkgadelekifesvknlskaakem 383

Qy         360 KNAVKELTS 368
               |:||||||
Db         384 tnsvkelts 392
```

```
RESULT 7
AAB62714
ID   AAB62714 standard; Protein; 369 AA.
XX
AC   AAB62714;
XX
DT   03-APR-2001  (first entry)
XX
DE   Borrelia sp chimeric ospc protein SEQ ID NO: 32.
XX
KW   Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
XX
PN   WO200078966-A1.
XX
PD   28-DEC-2000.
XX
PF   19-JUN-2000; 2000WO-US16915.
XX
PR   18-JUN-1999;  99US-0140042.
XX
PA   (UYNY ) UNIV NEW YORK STATE RES FOUND.
PA   (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI   Dattwyler RJ, Seinost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
PT   Compositions of OspC polypeptides from strains of Borrelia which cause
PT   Lyme disease -
XX
DR   WPI; 2001-050113/06.
DR   N-PSDB; AAF29018.
XX
PS   Claim 43; Page 89; 160pp; English.
XX
CC   The present invention provides compositions comprising ospc proteins and
CC   chimeric ospc proteins from members of the Borrelia genus. These may be
CC   Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC   vaccines against Borrelia infection, which is spread by ticks and leads
CC   to Lyme disease.
XX
SQ   Sequence  369 AA;

Query Match            82.3%;  Score 1479;  DB 22;  Length 369;
Best Local Similarity  83.8%;  Pred. No. 3e-87;
Matches  310;  Conservative   25;  Mismatches   31;  Indels    4;  Gaps    4;

Qy    2 ACNNSGKDGNASANSANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKAI 61
         |||||||||| ||||||||||||||||||||||| :|||||:||||||  :||||||| |||
Db    1 acnnsgkdgntsansadesvkgpnlteiskkitdsnavllavkeveallssidela-kai 59

Qy   62 GKKIGNNG-LEANQSKNTSLLSGAYAISDLIAEKLNVLKNEE-LKEKIDTAKQCSTEFTN 119
         ||||  :  | : ::|| |:|||| ||: | :||:|| |:||  |||| ||  :|| ||:|
Db   60 gkkikndgsldneanrnesllagaytistlltqklskingseglkeklaaakccseefst 119

Qy  120 KLKSEHAVLGLDNLTDDNAQRAILKKH-ANKDKGAAELEKLFKAVENLSKAAQDTLKNAV 178
         ||| :|| |:|   :::|| :: ||:  |  ||| ||||: :  |  |||| : |:||:|
Db  120 klkdnhaqlglggvtdenakkallkanaagkdkgveelekisgsleslskaakemlansv 179

Qy  179 KELTSPIVHGNNSRKDGNASTNSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASI 238
         |||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  180 keltspvvhgnnsrkdgnastnsadesvkgpnlteiskkitesnavvlavkevetllasi 239

Qy  239 DELATKAIGKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQC 298
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  240 delatkaigkkignngleanqskntsllsgayaisdliaeklnvlkneelkekidtakqc 299

Qy  299 STEFTNKLKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDT 358
         ||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db  300 steftnklksehavlgldnltddnaqrailkkhankdkgaaelekifkavenlskaaqdt 359

Qy  359 LKNAVKELTS 368
         ||||||||||
Db  360 lknavkelts 369
```

```
RESULT 8
AAB62732
ID   AAB62732 standard; Protein; 393 AA.
XX
AC   AAB62732;
XX
DT   03-APR-2001  (first entry)
XX
DE   Borrelia sp chimeric ospc protein SEQ ID NO: 68.
XX
KW   Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
XX
PN   WO200078966-A1.
XX
PD   28-DEC-2000.
XX
PF   19-JUN-2000; 2000WO-US16915.
XX
PR   18-JUN-1999;  99US-0140042.
XX
PA   (UYNY ) UNIV NEW YORK STATE RES FOUND.
PA   (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI   Dattwyler RJ, Seinost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
PT   Compositions of OspC polypeptides from strains of Borrelia which cause
PT   Lyme disease -
XX
DR   WPI; 2001-050113/06.
DR   N-PSDB; AAF29036.
XX
PS   Claim 133-134; 160pp; English.
XX
CC   The present invention provides compositions comprising ospc proteins and
CC   chimeric ospc proteins from members of the Borrelia genus. These may be
CC   Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC   vaccines against Borrelia infection, which is spread by ticks and leads
CC   to Lyme disease.
XX
SQ   Sequence  393 AA;

Query Match            82.1%;  Score 1476;  DB 22;  Length 393;
Best Local Similarity  83.5%;  Pred. No. 5.1e-87;
Matches  309;  Conservative   26;  Mismatches   31;  Indels    4;  Gaps    4;

Qy    2 ACNNSGKDGNASANSANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKAI 61
         :|||||||| ||||||||||||||||||||||| :|||||:||||||  :||||||| |||
Db   25 scnnsgkdgntsansadesvkgpnlteiskkitdsnavllavkeveallssidela-kai 83

Qy   62 GKKIGNNG-LEANQSKNTSLLSGAYAISDLIAEKLNVLKNEE-LKEKIDTAKQCSTEFTN 119
         ||||  :  | : ::|| |:|||| ||: | :||:|| |:||  |||| ||  :|| ||:|
Db   84 gkkikndgsldneanrnesllagaytistlltqklskingseglkeklaaakccseefst 143

Qy  120 KLKSEHAVLGLDNLTDDNAQRAILKKH-ANKDKGAAELEKLFKAVENLSKAAQDTLKNAV 178
         ||| :|| |:|   :::|| :: ||:  |  ||| ||||: :  |  |||| : |:||:|
Db  144 klkdnhaqlglggvtdenakkallkanaagkdkgveelekisgsleslskaakemlansv 203

Qy  179 KELTSPIVHGNNSRKDGNASTNSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASI 238
         |||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  204 keltspvvhgnnsrkdgnastnsadesvkgpnlteiskkitesnavvlavkevetllas 263
```

```
Qy    239  DELATKAIGKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQC  298
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    264  delatkaigkkignngleanqskntsllsgayaisdliaeklnvlkneelkekidtakqc  323

Qy    299  STEFTNKLKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDT  358
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    324  steftnklksehavlgldnltddnaqrailkkhankdkgaaeleklfkavenlskaaqdt  383

Qy    359  LKNAVKELTS  368
           ||||||||||
Db    384  lknavkelts  393


RESULT  9
AAB62713
ID   AAB62713 standard; Protein; 377 AA.
XX
AC   AAB62713;
XX
DT   03-APR-2001   (first entry)
XX
DE   Borrelia sp chimeric ospC protein SEQ ID NO: 30.
XX
KW   Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
XX
PN   WO200078966-A1.
XX
PD   28-DEC-2000.
XX
PF   19-JUN-2000; 2000WO-US16915.
XX
PR   18-JUN-1999; 99US-0140042.
XX
PA   (UINV )  UNIV NEW YORK STATE RES FOUND.
PA   (BROO-)  BROOK BIOTECHNOLOGIES INC.
XX
PI   Dattwyler RJ, Seinost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
PS   Claim 43; Page 86-87; 160pp; English.
XX
PT   The present invention provides compositions comprising ospC proteins and
PT   chimeric ospC proteins from members of the Borrelia genus. These may be
PT   Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
PT   vaccines against Borrelia infection, which is spread by ticks and leads
PT   to Lyme disease.
XX
DR   WPI; 2001-050113/06.
DR   N-PSDB; AAF29017.
XX
CC   The present invention provides compositions comprising ospC proteins and
CC   chimeric ospC proteins from members of the Borrelia genus. These may be
CC   Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC   vaccines against Borrelia infection, which is spread by ticks and leads
CC   to Lyme disease.
SQ   Sequence  377 AA;

Query Match             68.2%;  Score 1225.5;  DB 22;  Length 377;
Best Local Similarity   69.5%;  Pred. No. 4.9e-71;
Matches  262;  Conservative   43;  Mismatches   61;  Indels   11;  Gaps   4;

Qy      1  MACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKA   60
           |||||||||||| |||||||||||||||| ||| |||||| ||||||| ||:||||||| ||
Db      1  macnnsgkdgntsansadesvkgpnlteinkkitdsnavllavkeveallssidelaaka   60

Qy     61  IGKKI-GNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTN  119
           ||||  :|||||||||||||||||||||||||||||||||||  :||||||||||||||
Db     61  igkkihqnngldtennhngsllagayaistlikqkldglkneglkekidaakccsetftn  120

Qy    120  KLKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVK  179
```

```
Db    121  klkekhtdlgkegvtdadakeallkangtktkgaaelgklfesvevlskaakemlansvk  180
           |||  :|   ::|||:  || || ::|||||||:|||||||::||||  |||||  :||
Qy    180  ELISPIVHG----------NNSRKDGNASTNSADESVKGPNLTEISKKITESNAVVLAVKEV  231
           |||:|:|:||          |||||||||||||||||||||||||||||||||||||:
Db    181  eltspvvaespkkpsmvnnsgkdgntsansadesvkgpnlteiskkitesnavvlavkei  240

Qy    232  ETLLASIDELATKAIGKKI-GNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKN-EELK  289
           ::||||||||||||||||| |||| ||||||||:|||||||||:|||||:|||||: ||||
Db    241  etllasidelatkaigkkiqgnngglaveaghngtllagaytiskltqkldglknsekl k  300

Qy    290  EKIDTAKQCSTEFTNKLKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVE  349
           ||:| |||  :| ::|||::||| |||:|||:|||::|| ||||||||||||| |||||||
Db    301  ekienakkcsedfkkleqehaqlgienvtdenakkalitdaakdkgaaelgklfave    360

Qy    350  NLSKAAQDTLKNAVKEL  366
           ||:||||  ||||||||
Db    361  nlakaakemlansvkel  377


RESULT  10
AAB62711
ID   AAB62711 standard; Protein; 373 AA.
XX
AC   AAB62711;
XX
DT   03-APR-2001   (first entry)
XX
DE   Borrelia sp chimeric ospC protein SEQ ID NO: 26.
XX
KW   Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
XX
PN   WO200078966-A1.
XX
PD   28-DEC-2000.
XX
PF   19-JUN-2000; 2000WO-US16915.
XX
PR   18-JUN-1999; 99US-0140042.
XX
PA   (UINV )  UNIV NEW YORK STATE RES FOUND.
PA   (BROO-)  BROOK BIOTECHNOLOGIES INC.
XX
PI   Dattwyler RJ, Seinost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
PS   Claim 43; Page 81; 160pp; English.
XX
PT   Compositions of OspC polypeptides from strains of Borrelia which cause
PT   Lyme disease are used to immunize animals and detect immune responses
PT   to Lyme disease.
XX
DR   WPI; 2001-050113/06.
DR   N-PSDB; AAF29015.
XX
CC   Compositions of OspC polypeptides from strains of Borrelia which cause
CC   Lyme disease are used to immunize animals and detect immune responses
CC   to Lyme disease.
SQ   Sequence  373 AA;

Query Match             68.1%;  Score 1224;  DB 22;  Length 373;
Best Local Similarity   70.2%;  Pred. No. 6e-71;
Matches  262;  Conservative   42;  Mismatches   61;  Indels    8;  Gaps   4;

Qy      2  ACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKAI   61
           |||||||||| |||||||||||||||||||| |||:|||| ||||||| ||:|||||||||
Db      1  acnnsgkdgntsansadesvkgpnlteiskkitdsnavllavkeveallssidelaakai   60
```

```
Qy    62  GKKI-GNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNK  120
              ||||  ||||||  |  ||  ||||||  ||||  |  |||||||  ||||||||  ||||
Db    61  gkkihgnngldteynhngsllagayaistlikqkldglkneglkekidaakkcsetftnk  120

Qy   121  LKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKE  180
              ||:||||||  |  ||  ||:||||||  ||||  ||  ::||  :  |  |  |  |||
Db   121  lkekhtdigkegvtdadakeailktngtktkgaeelgklfesveviskaakemlansvke  180

Qy   181  LTSPIVHG-----NNSRKDGNASTNSADESVKGPNLTEISKKITESNAVVLAVKEVETLL  235
              ||:||  |      ||:|||||  ||||||||  ||  ||  :::|||||||  ||  |
Db   181  ltspvvaespamvnnsgkdgntsansadesvkgpnlteiskkitesnavvlavkeietll  235

Qy   236  ASIDELATKAIGKKI-GNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKN-EELKEKID  293
              ||||||||||||||  ||||||||||||||||||||||||||||||||||  |||||||
Db   241  asidelatkaigkkiqgngglaveaghngtllagaytiskiltqkidglknseklkekie  300

Qy   294  TAKQCSTEFTNKLKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSK  353
              |||  |  ||  ||||||||||||||||||||||||||||||||||||||||||||||
Db   301  nakkcsedftkklegehaqlgienvcdenakkailitdaakdkgaaeleklfkavenlak  360

Qy   354  AAQDTLKNAVKEL  366
              ||  |  :  |||
Db   361  aakemlansvkel  373
```

```
RESULT 11
ID   AAB62729
AAB62729 standard; Protein; 397 AA.
XX
AC   AAB62729;
XX
DT   03-APR-2001  (first entry)
XX
DE   Borrelia sp chimeric ospc protein SEQ ID NO: 62.
XX
KW   Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
XX
XX
PN   WO200078966-A1.
XX
PD   28-DEC-2000.
XX
PF   19-JUN-2000; 2000WO-US16915.
XX
PR   18-JUN-1999;  99US-0140042.
XX
PA   (UYNY ) UNIV NEW YORK STATE RES FOUND.
PA   (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI   Dattwyler RJ, Selnost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
DR   WPI; 2001-050113/06.
DR   N-PSDB; AAF29033.
XX
PT   Compositions of OspC polypeptides from strains of Borrelia which cause
PT   to Lyme disease -
XX
PS   Claim 43; Page 125-126; 160pp; English.
XX
CC   The present invention provides compositions comprising ospc proteins and
CC   chimeric ospC proteins from members of the Borrelia genus. These may be
CC   Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC   vaccines against Borrelia infection, which is spread by ticks and leads
CC   to Lyme disease.
XX
SQ   Sequence  397 AA;
```

```
Query Match       67.9%;  Score 1221;  DB 22;  Length 397;
```

```
Best Local Similarity  70.0%;  Pred. No. 1e-70;
Matches  261;  Conservative  43;  Mismatches  61;  Indels  8;  Gaps  4;
```

```
Qy     2  ACNNSGKDGNSASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKAI   61
             ||  |||| |  |||||||  ||||||||||||||||||||||||||||||  ||:|||||
Db    25  scnnsgkdgntsansadesvkgpnlteiskkitdsnavllavkeveallssidelaakal   84

Qy    62  GKKI-GNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNK  120
             ||||  |||||  |  |  ||  ||||||  ||  |  |||||||  ||||||||  ||||
Db    85  gkkihgnngldteynhngsllagayaistlikqkldglkneglkekidaakkcsetftnk  144

Qy   121  LKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKE  180
             ||:||||||  |  ||  ||:||||||  ||||  ||  ::||  :  |  |  |  |||
Db   145  lkekhtdigkegvtdadakeailktngtktkgaeelgklfesveviskaakemlansvke  204

Qy   181  LTSPIVHG-----NNSRKDGNASTNSADESVKGPNLTEISKKITESNAVVLAVKEVETLL  235
             ||:||  |      ||:|||||  ||||||||  ||  ||  :::|||||||  ||  |
Db   205  ltspvvaespamvnnsgkdgntsansadesvkgpnlteiskkitesnavvlavkeietll  264

Qy   236  ASIDELATKAIGKKI-GNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKN-EELKEKID  293
             ||||||||||||||  ||||||||||||||||||||||||||||||||||  |||||||
Db   265  asidelatkaigkkiqgngglaveaghngtllagaytiskiltqkidglknseklkekie  324

Qy   294  TAKQCSTEFTNKLKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSK  353
             |||  |  ||  |||||||||||||||||||||||||||||||||||||||||||||||
Db   325  nakkcsedftkklegehaqlgienvcdenakkailitdaakdkgaaeleklfkavenlak  384

Qy   354  AAQDTLKNAVKEL  366
             ||  |  :  |||
Db   385  aakemlansvkel  397
```

```
RESULT 12
ID   AAB62716
AAB62716 standard; Protein; 369 AA.
XX
AC   AAB62716;
XX
DT   03-APR-2001  (first entry)
XX
DE   Borrelia sp chimeric ospc protein SEQ ID NO: 36.
XX
KW   Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
XX
XX
PN   WO200078966-A1.
XX
PD   28-DEC-2000.
XX
PF   19-JUN-2000; 2000WO-US16915.
XX
PR   18-JUN-1999;  99US-0140042.
XX
PA   (UYNY ) UNIV NEW YORK STATE RES FOUND.
PA   (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI   Dattwyler RJ, Selnost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
DR   WPI; 2001-050113/06.
DR   N-PSDB; AAF29020.
XX
PT   Compositions of OspC polypeptides from strains of Borrelia which cause
PT   to Lyme disease -
XX
PS   Claim 43; Page 94-95; 160pp; English.
XX
CC   The present invention provides compositions comprising ospc proteins and
CC   chimeric ospC proteins from members of the Borrelia genus. These may be
CC   Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC   vaccines against Borrelia infection, which is spread by ticks and leads
```

Left column

```
CC   to Lyme disease.
XX
SQ   Sequence   369 AA;

Query Match        67.9%;  Score 1220;  DB 22;  Length 369;
Best Local Similarity  70.5%;  Pred. No. 1.1e-70;
Matches 261;  Conservative 45;  Mismatches 58;  Indels 6;  Gaps 6;

Qy    2 ACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKAI 61
        ||||||| ||| ||||||||||||||||||||| ||||:||||||||| ||||||||| ||
Db    1 acnnsgkdgntsansadesvkgpnlteiskkitdsnavllavkeveallssidela-kai 59

Qy   62 GKKIGNNG-LEANQSKNTSLLSGAYAISDLIAEKLNVLKNEE-LIKEKIDTAKQCSTEFTN 119
        |||| || : || :: :|||||||||||||:|| | |: | | : | |:| |||
Db   60 gkkikndgsldneanrnesllagaytistliitqkiskingsseglkekiaaakkcseefst 119

Qy  120 KLKSEHAVLGLDNLITDDNAQRAILKKH-ANKDKGAAAELEKLFKAVENLSKAAQDTLKNAV 178
        |||  || :| |: | |:  :|||:  | :|| :|||  |||| || ||: | |||||| |
Db  120 klkdnhaqlgiqgyvdenakkailkanaagkdkgveeleklsgsleslskaakemlansv 179

Qy  179 KELTSPIVHGNNSRKDGNASTNSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASI 238
        ||||||:|||||| |||| ||||||||||||||||||||||||||||||||||| |||||
Db  180 keltspvvhgnnsgkdgntsansadesvkgpnlteiskkitesnavllavkeietllasi 239

Qy  239 DELATKAIGKKI-GNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKN-EELKEKIDTAK 296
        ||||||||||| :||||:  | || :|||||||:|| | :|:|| || ||:||||:||
Db  240 delatkaigkkigqnggllaveaghngtllagaytiskiltqkidglknsekikekienak 299

Qy  297 QCSTEFTNKLKSEHAVLGLDNLITDDNAQRAILKKHANKDKGAAAELEKLFKAVENLSKAAQ 356
        :| |:| |||  || :| |:  | |:  :|||| |:|||||| |||||||||||: | ||
Db  300 kcsedftkklegehaqlgienvtdenakkailtdaakdkgaaeleklfkavenlakaak 359

Qy  357 DTLKNAVKEL 366
        : | :|:|||
Db  360 emlansvkel 369


RESULT 13
AAB62739
ID   AAB62739 standard; Protein; 400 AA.
XX
AC   AAB62739;
XX
DT   03-APR-2001  (first entry)
XX
DE   Borrelia sp chimeric ospc protein SEQ ID NO: 82.
XX
KW   Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
XX
PN   WO200078966-A1.
XX
PD   28-DEC-2000.
XX
PF   19-JUN-2000;  2000WO-US16915.
XX
PR   18-JUN-1999;  99US-0140042.
XX
PA   (UYNY ) UNIV NEW YORK STATE RES FOUND.
PA   (BROO-) BROOK BIOTECHNOLOGIES INC.
PI   Dattwyler RJ, Seinost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
DR   WPI; 2001-050113/06.
DR   N-PSDB; AAF29043.
XX
PT   Compositions of OspC polypeptides from strains of Borrelia which cause
PT   Lyme disease are used to immunize animals and detect immune responses
PT   to Lyme disease -
```

Right column

```
PS   Claim 43; Page 153; 160pp; English.
XX
CC   The present invention provides compositions comprising ospC proteins and
CC   chimeric ospc proteins from members of the Borrelia genus. These may be
CC   chimeric ospC proteins from members of the Borrelia genus. These may be
CC   Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC   vaccines against Borrelia infection, which is spread by ticks and leads
CC   to Lyme disease.
XX
SQ   Sequence   400 AA;

Query Match        67.9%;  Score 1219.5;  DB 22;  Length 400;
Best Local Similarity  69.4%;  Pred. No. 1.3e-70;
Matches 261;  Conservative 43;  Mismatches 61;  Indels 11;  Gaps 4;

Qy    2 ACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKAI 61
        ||||||| ||| |||||||||||||||||||||| ||||:|||||||||||||||| |||
Db   25 scnnsgkdgntsansadesvkgpnlteiskkitdsnavllavkeveallssidelaakal 84

Qy   62 GKKI-GNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNK 120
        ||||  |:| | :|: :|||||||||:|| | :|:|| ||||||||:| ||| ||||||
Db   85 gkkihgnngldteynhngsllagayaistliikgkldglkneglkekidaakkcsetftnk 144

Qy  121 LKSEHAVLGLDNLITDDNAQRAILKKHANKDKGAAAELEKLFKAVENLSKAAQDTLKNAVKE 180
        || | || :|  : | :  :|||| |:|:|:||||| |||| |: || ||| | |||||||
Db  145 lkekhtdlgkegytdadakeailktngtktkgaaelgklfesvevlskaakemlansvke 204

Qy  181 LTSPIVHG------NNSRKDGNASTNSADESVKGPNLTEISKKITESNAVVLAVKEVE 232
        ||||:|||      ||| |||| ||||||||||||||||||||||||||||||||||
Db  205 ltspvvaespkkpswnnsgkdgntsansadesvkgpnlteiskkitesnavlavkele 264

Qy  233 TLLASIDELATKAIGKKI-GNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKN-EELKE 290
        |||||||||||||||||| :||||:  | || :|||||||:|| | :|:|| || ||:||
Db  265 tllasidelatkaigkkigqnggllaveaghngtllagaytiskiltqkidglknsekke 324

Qy  291 KIDTAKQCSTEFTNKLKSEHAVLGLDNLITDDNAQRAILKKHANKDKGAAAELEKLFKAVEN 350
        ||:||:|| |:| |||| | || :|  : | :  :|||| |:|||||| |||||||||||
Db  325 kienakkcsedftkklegehaqlgienvtdenakkailtdaakdkgaaeleklfkaven 384

Qy  351 LSKAAQDTLKNAVKEL 366
        |:|| | |:|:|:|||
Db  385 lakaakemlansvkel 400


RESULT 14
AAB62731
ID   AAB62731 standard; Protein; 393 AA.
XX
AC   AAB62731;
XX
DT   03-APR-2001  (first entry)
XX
DE   Borrelia sp chimeric ospc protein SEQ ID NO: 66.
XX
KW   Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
XX
PN   WO200078966-A1.
XX
PD   28-DEC-2000.
XX
PF   19-JUN-2000;  2000WO-US16915.
XX
PR   18-JUN-1999;  99US-0140042.
XX
PA   (UYNY ) UNIV NEW YORK STATE RES FOUND.
PA   (BROO-) BROOK BIOTECHNOLOGIES INC.
PI   Dattwyler RJ, Seinost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
DR   WPI; 2001-050113/06.
```

DR   N-PSDB; AAF29035.
XX
PA   (unreadable)
PT   Compositions of OspC polypeptides from strains of Borrelia which cause
PT   Lyme disease are used to immunize animals and detect immune responses
PT   to Lyme disease -
XX
PT   Lyme disease.
XX
PS   Claim 43; Page 131; 160pp; English.
XX
CC   The present invention provides compositions comprising ospc proteins and
CC   chimeric ospC proteins from members of the Borrelia genus. These may be
CC   Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC   vaccines against Borrelia infection, which is spread by ticks and leads
CC   to Lyme disease.
XX
SQ   Sequence 393 AA;

Query Match 67.7%; Score 1217; DB 22; Length 393;
Best Local Similarity 70.3%; Pred. No. 1.8e-70;
Matches 260; Conservative 46; Mismatches 58; Indels 6; Gaps 6;

```
Qy    2  ACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKAI   61
            ||||||||||||| |||||||||||||||||: ||||||||||||||||| || ||
Db   25  scnnsgkdgntsansadesvkgpnlteiskkitdsnavllavkeveallssidela-kai   83

Qy   62  GKKIGNNG-LEANQSKNTSLLSGAYAISDLIAEKLNVLKNEE-LKEKIDTAKQCSTEFIN  119
            |||||||  |||||||||||||||:||||||||||||||  ||:| |||||||| ||
Db   84  gkkikndgsldneanrnesllagaytistiltqkislngseglkekiaaakkcseefst  143

Qy  120  KLKSEHAVLGLDNLTDDNAQRAILKKH-ANKDKGAAELEKLFKAVENLSKAAQDTLKNAV  178
            :||:||||||||||||||||||::||: |||||||:||||||||||:|||||: |:|
Db  144  klkdnhaqlgigvtdenakkallkanaagkdkgveelekisgslesiskaakemlansv  203

Qy  179  KELTSPIVHGNNSRKDGNASTNSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASI  238
            |||:||:| ||:|:|||||:|||||||||||||||||||||||||||||||:|||||||
Db  204  keltspvvhgnnsgkdgntsansadesvkgpnlteiskkitesnavvlavkeietllasi  263

Qy  239  DELATKAIGKKI-GNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKN-EELKEKIDTAK  296
            ||||||||:||| ||||:|||||||||||||||||||||||||||||: |||||||||||
Db  264  delatkaigkkiqgngglaveaghngtilagaytislkiidglknseklkeklenak    323

Qy  297  QCSTEFTNKLKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQ  356
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db  324  kcsedfctklegehaqigienvcdenakkalltdaakdkgaaeleklfkavenlakaak  383

Qy  357  DTLKNAVKEL  366
            :|:|||||||
Db  384  emlansvkel  393
```

RESULT 15
AAB62712
ID   AAB62712 standard; Protein; 378 AA.
XX
AC   AAB62712;
XX
DT   03-APR-2001 (first entry)
XX
DE   Borrelia sp chimeric ospc protein SEQ ID NO: 28.
XX
KW   Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
XX
PN   WO200078966-A1.
XX
PD   28-DEC-2000.
XX
PF   19-JUN-2000; 2000WO-US16915.
XX
PR   18-JUN-1999; 99US-0140042.

PA   (UYNY) UNIV NEW YORK STATE RES FOUND.
PA   (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI   Dattwyler RJ, Seinost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
DR   WPI; 2001-050113/06.
DR   N-PSDB; AAF29016.
XX
PT   Compositions of OspC polypeptides from strains of Borrelia which cause
PT   Lyme disease are used to immunize animals and detect immune responses
PT   to Lyme disease -
XX
PS   Claim 43; Page 83-84; 160pp; English.
XX
CC   The present invention provides compositions comprising ospc proteins and
CC   chimeric ospC proteins from members of the Borrelia genus. These may be
CC   Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC   vaccines against Borrelia infection, which is spread by ticks and leads
CC   to Lyme disease.
XX
SQ   Sequence 378 AA;

Query Match 65.6%; Score 1178; DB 22; Length 378;
Best Local Similarity 67.3%; Pred. No. 5.3e-68;
Matches 255; Conservative 47; Mismatches 65; Indels 12; Gaps 5;

```
Qy    1  MACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKA   60
            ||||||||||||:||||||||||||||||||||||:||||||||||||||||:|||:||
Db    1  macnnsgkdgntsansadesvkgpnlteiskkitdsnavllavkeveallssidelaaka   60

Qy   61  IGKKI-GNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTN  119
            |||| |||||:|||||||||||||||||||||:|||:|| ::||||||||||||:||
Db   61  igkkihqgngldtenhhngsllagayaistilkqkidglkneglkekidaakkcsetftn  120

Qy  120  KLKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVK  179
            :||||||:||||||||||:||||||||:||||:||||||||:|||:|||||:||||||
Db  121  klkekhtdlgkegvtdadakealikangtktkqaaelgklfesvevlskaakemlansvk  180

Qy  180  ELTSPIVHG------NNSRKDGNASTNSADESVKGPNLTEISKKITESNAVVLAVKEV   231
            ||:|| :  ||:||||||:||||||||||||||||||||||||||||||||||||
Db  181  eltspvvaespkkpsmvnnsgkdgntsansadesvkgpnlteiskkitesnavvlavkev  240

Qy  232  ETLLASIDELATKAIGKKIGNN-GLEANQSKNTSLLSGAYAISDLIAEKLNVLKNE-ELK  289
            :|||:|:|||| :|||||| |||  |||||||||||||||||||||||||||||:| |||
Db  241  etlltsidela-kaigkkikndvslsidneadhngslisgaylisnlitkkisaikdsgelk  299

Qy  290  EKIDTAKQCSTEFTNKLKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVE  349
            ||||||:||||||:|||||||||||||||||||:||:||||||||:||||||||||||:
Db  300  aeiekkkcseeftaklkgehtdlgkegvtdadnakkaliktnndktkgadeleklfesvk  359

Qy  350  NLSKAAQDTLKNAVKELTS  368
            |||||||:||:|:|:||:|
Db  360  nlskaakemltnsvkelts  378
```

Search completed: March 18, 2002, 09:54:33
Job time: 331 sec

OM protein - protein search, using sw model

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

Run on: March 18, 2002, 09:55:34 ; Search time 55.5 Seconds
(without alignments)
149.211 Million cell updates/sec

Title: US-09-596-746A-38
Perfect score: 1797
Sequence: 1 MACNNSGKDGNASANSADES................ENLSKAAQDTLKNAVKELTS 368

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

arched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	673	37.5	466	4	US-08-235-836C-110	Sequence 110, App
2	657.5	36.6	212	1	US-08-158-353-2	Sequence 2, Appli
3	640	35.6	588	4	US-08-235-836C-122	Sequence 122, App
4	638.5	35.5	466	4	US-08-235-836C-107	Sequence 107, App
5	623.5	34.7	209	4	US-09-196-293-15	Sequence 15, Appl
6	623.5	34.7	210	1	US-08-158-353-3	Sequence 3, Appl
7	617.5	34.4	210	4	US-08-235-836C-30	Sequence 30, Appl
8	616.5	34.3	210	4	US-09-196-293-11	Sequence 11, Appl
9	588.5	32.7	212	1	US-08-158-353-4	Sequence 4, Appl
10	588.5	32.7	212	4	US-08-209-603E-15	Sequence 15, Appl
11	584.5	32.5	212	1	US-08-031-295-2	Sequence 2, Appl
12	584.5	32.5	212	5	PCT-US93-07261-16	Sequence 16, Appl
13	583.5	32.5	212	5	US-08-235-836C-34	Sequence 34, Appl
14	578.5	32.2	212	4	US-07-903-580-2	Sequence 2, Appl
15	578.5	32.2	212	4	US-09-196-293-11	Sequence 11, Appl
16	563	31.3	207	4	US-08-209-603E-11	Sequence 11, Appl
17	536	29.8	209	4	US-08-235-836C-36	Sequence 36, Appl
18	157.5	8.8	1588	5	US-08-235-836C-32	Sequence 32, Appl
19	157.5	8.8	1663	5	PCT-US93-07261-16	Sequence 16, Appl
20	150	8.3	695	1	US-08-470-260-5	Sequence 5, Appl
21	150	8.3	695	1	US-08-482-847-23	Sequence 23, Appl
22	150	8.3	1147	3	US-08-471-491-5	Sequence 5, Appl
23	150	8.3	1147	3	US-08-466-662-5	Sequence 5, Appl
24	150	8.3	3289	2	US-08-477-451-2	Sequence 2, Appl
25	149.5	8.3	1147	2	US-08-144-121-4	Sequence 4, Appl
26	149.5	8.3	1196	1	US-08-735-893-4	Sequence 4, Appl
27	149	8.3	630	4	US-08-973-462-9	Sequence 9, Appl
28	141.5	7.9	1786	4	US-08-973-462-8	Sequence 8, Appl
29	141	7.8	859	1	US-08-053-614-2	Sequence 2, Appl
30	141	7.8	859	1	US-08-316-397B-2	Sequence 2, Appl
31	141	7.8	859	1	US-09-034-306-2	Sequence 2, Appl
32	141	7.8	859	4	US-09-259-437-2	Sequence 2, Appl
33	141	7.8	859	5	PCT-US93-09782-2	Sequence 2, Appl
34	141	7.8	1181	1	US-08-053-614-4	Sequence 4, Appl
35	141	7.8	1181	1	US-08-316-397B-4	Sequence 4, Appl
36	141	7.8	1181	2	US-09-034-306-4	Sequence 4, Appl
37	141	7.8	1181	4	US-09-259-437-4	Sequence 4, Appl
38	141	7.8	1181	5	PCT-US93-09782-4	Sequence 6, Appl
39	139.5	7.8	2482	1	US-08-328-254-6	Sequence 6, Appl
40	138.5	7.7	619	1	US-08-465-746-2	Sequence 2, Appl
41	138.5	7.7	619	1	US-08-214-164-2	Sequence 2, Appl
42	138.5	7.7	619	2	US-08-467-852A-3	Sequence 3, Appl
43	138.5	7.7	619	2	US-08-246-636-2	Sequence 3, Appl
44	138.5	7.7	619	2	US-08-247-491A-3	Sequence 3, Appl
45	138.5	7.7	619	2	US-08-319-795-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-08-235-836C-110
Sequence 110, Application US/08255836C

Patent No. 6248562
GENERAL INFORMATION:
APPLICANT: Dunn, John J.
APPLICANT: Luft, Benjamin J.
TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising
TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor
NUMBER OF SEQUENCES: 144
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brookhaven National Laboratory
STREET:
CITY: Upton
STATE: NY
COUNTRY: USA
ZIP: 11973
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/235,836C
FILING DATE: 29-APR-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/148,191
FILING DATE: 01-11-93
ATTORNEY/AGENT INFORMATION:
NAME: Bogosian, Margaret C.
REGISTRATION NUMBER: 25,324
REFERENCE/DOCKET NUMBER: BNL93-28A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 282-7338
TELEFAX: (516) 282-3729
INFORMATION FOR SEQ ID NO: 110:
SEQUENCE CHARACTERISTICS:
LENGTH: 466 amino acid
TYPE: amino acids
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-235-836C-110

Qy

Query Match 37.5%; Score 673; DB 4; Length 466;
Best Local Similarity 44.9%; Pred. No. 1.3e-45;
Matches 173; Conservative 58; Mismatches 112; Indels 42; Gaps 11;

Qy 1 MACNNSGKDGNASANSADESVKGPNLTEISRKITESNAVVLAVKEVETLLASIDELATKA 60

RESULT 2
US-08-158-353-2
Sequence 2, Application US/08158353
Patent No. 5620662
GENERAL INFORMATION:
APPLICANT: Padula, Steven J.
TITLE OF INVENTION: Methods for Diagnosing Early Lyme
TITLE OF INVENTION: Disease
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Militia Drive
CITY: Lexington
STATE: MA
COUNTRY: USA
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/158,353
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Carroll, Alice O.
REGISTRATION NUMBER: 33,542
REFERENCE/DOCKET NUMBER: UCT93-05
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-861-6240
TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 212 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-158-353-2

Query Match   36.6%;   Score 657.5;   DB 1;   Length 212;
Best Local Similarity   72.9%;   Pred. No. 7.6e-45;
Matches 140;   Conservative 20;   Mismatches 29;   Indels 3;   Gaps 2;

Qy   17   ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVILAVKEVEALLSSIDEIAAKA   76
Db   17   IGKKI-GNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTN   61
Qy   77   IGKKIHQNNGLDTEVNHNGSLLAGRYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN   136
Db   120   KLKSEHAVIGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVK   179
Qy   137   KLKAKHTDLGKEGYTDADAKEAILKTNGTRTKGAEELGKLFESEVLSKAAKEMLANSVK   196
Db   180   ELTSPIVHGNNSRKDGNAST----NSADESVKGPNLTEISKKITESNAVVLAVKEVETLL   235
Qy   197   ELTSPVVAESPKFPKQNVSSLDEKNSVSVDLPGEMKVLVSKEKNKDGKY------DLI   248
Db   236   ASIDELATRAIGKKIGNN------GLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEE--   287
Qy   249   ATVDKLELKGTSDK--NNGSGVLEGVKADKSKVKLTIS-----DDLGQTTLEVFKEDGKT   301
Db   288   -LKEKIDTAKQCSTE--FTNKLKSEHAVIGLDNLTDDNAQRAILKKHANKDKGAAELEKL   344
Qy   302   LVSKKVTSKDKSSTEEKFNEKGEVSEKII------TRADGTRLEYTGIKSDGSGKAKEV   354
Db   345   FK--AVENLSKAAQDTLKNAVKELT   367
Qy   355   LKGYVLEGTLITAEKTTL--VVKEGT   377

RESULT 3
US-08-235-836C-122
Sequence 122, Application US/08235836C
Patent No. 6248562
GENERAL INFORMATION:
APPLICANT: Dunn, John J.
APPLICANT: Luft, Benjamin J.
TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor
TITLE OF INVENTION: No. 6248562e1 Chimeric Proteins Comprising
NUMBER OF SEQUENCES: 144
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brookhaven National Laboratory
STREET: Upton
CITY:
STATE: NY
COUNTRY: USA
ZIP: 11973
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/235,836C
FILING DATE: 29-APR-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/148,191
FILING DATE: 01-11-93
ATTORNEY/AGENT INFORMATION:
NAME: Bogosian, Margaret C.
REGISTRATION NUMBER: 25,324
REFERENCE/DOCKET NUMBER: BNL93-28A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 282-7338
TELEFAX: (516) 282-3729
INFORMATION FOR SEQ ID NO: 122:
SEQUENCE CHARACTERISTICS:
LENGTH: 588 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-235-836C-122

Query Match   35.6%;   Score 640;   DB 4;   Length 588;
Best Local Similarity   39.1%;   Pred. No. 6.8e-43;
Matches 184;   Conservative 55;   Mismatches 110;   Indels 122;   Gaps 17;

Qy   1   MACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKA   60
Db   17   ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVILAVKEVEALLSSIDEIAAKA   76
Qy   61   IGKKIGNGNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKN-EELKEKIDTAKQCSTEFTN   119
Db   17   IGNLIAQNGLNAGANQNGSLLAGAYVISTLIAEKLDGLKNSEELKEKIEDAKKCNKAFTD   136
Qy   120   KLKSEHAVIGLDN--LTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNA   177
Db   137   KLKSSHAELGIANGAATDANAKAAILKTNGTKDKGAQELEKLFESEVKNLSKAAQETLNNS   196
Qy   178   VKELTSPIVHGN 189
Db   197   VKELTSPVVAEN 208

Qy   7   GKDGNASANSADESVKG--------PNLTEISKKITESNAVVLAVKEVET----LLASID   54
Db   9   GSQKENDLNEEDSSKKSHQNAKQDLPAVTEDSVSLFNGNKIFVSKEKNSSGKYDLRATID   68
Qy   55   ELATKAIGKKIGNNG---LEANQ-SKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTA   110

RESULT 4
US-08-235-836C-107
Sequence 107, Application US/0823583GC
Patent No. 6248562
GENERAL INFORMATION:
   APPLICANT: Dunn, John J.
   APPLICANT: Luft, Benjamin J.
   TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising
   TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor
   NUMBER OF SEQUENCES: 144
   CORRESPONDENCE ADDRESS:
      ADDRESSEE: Brookhaven National Laboratory
      STREET:
      CITY: Upton
      STATE: NY
      COUNTRY: USA
      ZIP: 11973
   COMPUTER READABLE FORM:
      MEDIUM TYPE: Floppy disk
      COMPUTER: IBM PC compatible
      OPERATING SYSTEM: PC-DOS/MS-DOS
      SOFTWARE: PatentIn Release #1.0, Version #1.25
   CURRENT APPLICATION DATA:
      APPLICATION NUMBER: US 08/235,836C
      FILING DATE: 29-APR-1994
      CLASSIFICATION: 435
   PRIOR APPLICATION DATA:
      APPLICATION NUMBER: US 08/148,191
      FILING DATE: 01-11-93
   ATTORNEY/AGENT INFORMATION:
      NAME: Bogosian, Margaret C.
      REGISTRATION NUMBER: 25,324
      REFERENCE/DOCKET NUMBER: BNL93-28A
   TELECOMMUNICATION INFORMATION:
      TELEPHONE: (516) 282-7338
      TELEFAX: (516) 282-3729
   INFORMATION FOR SEQ ID NO: 107:
      SEQUENCE CHARACTERISTICS:
         LENGTH: 466 amino acids
         TYPE: amino acid
         TOPOLOGY: linear
         MOLECULE TYPE: protein
US-08-235-836C-107

Query Match   35.5%;   Score 638.5;   DB 4;   Length 466;

```
Db   69 QVELKGTSDK--NNGSGTLEGSKPDKSKVKLTVSADLNTVTLEAFDA-SNQKISSKV-TK 124
Qy  111 KQCS-TEFT--NKL-----KSEHAVLGLDNLITD-DNAQRAI-------------TK 142
Db  125 KQGSITEETLKANKLDSKKLTRSNGTTLEYSQITDADNATKAVETLKNSIKLEGSLVVGK 184
Qy  143 ----------------------LKKHANKD---------------------- 150
Db  185 TTVEIKEGTVTLKREIEKDGKVKVFLNDTAGSNKKTGKWEDSTSTLTISADSKKTKDLVF 244
Qy  151 --KGAAELEKLFKAVENLSKAAQDTLKNAVKELTSPIVHGNNSRKDGNAS 198
Db  245 LTDGTITVQQYNTAGTSLEGSASEIKNLSE---LKNALK--GHPM--GNNSGKDGNTS 295
Qy  199 TNSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKAIGKKI-GNNGLEA 257
Db  296 ANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGKKIHQNNGLDT 355
Qy  258 NQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNKLKSEHAVLGLDN 317
Db  356 EYNHNGSLLAGRYAISTLIKQKLDGLIKNEGLKEKIDAAKKCSETFTNKLKEKHTDLGKEG 415
Qy  318 LTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKELTS 368
Db  416 VTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKELTS 466
```

RESULT 5
US-09-196-293-15
Sequence 15, Application US/0919293
Patent No. 6183755
GENERAL INFORMATION:
   APPLICANT: Fuchs, Renate
   APPLICANT: Motz, Manfred
   APPLICANT: Soutscheck, Erwin
   APPLICANT: Wilske, Bettina
   APPLICANT: Preac-Mursic, Vera
   TITLE OF INVENTION: Active proteins from Borrelia
   TITLE OF INVENTION: burgdorferi
   FILE REFERENCE: 738.001US2
   CURRENT APPLICATION NUMBER: US/09/196,293
   CURRENT FILING DATE: 1998-11-19
   EARLIER APPLICATION NUMBER: US 08/209,603
   EARLIER FILING DATE: 1994-03-10
   EARLIER APPLICATION NUMBER: US 07/862,535
   EARLIER FILING DATE: 1992-06-19
   EARLIER APPLICATION NUMBER: WO PCT/EP90/02282
   EARLIER FILING DATE: 1990-12-21
   EARLIER APPLICATION NUMBER: DE P39 42 728.5
   EARLIER FILING DATE: 1989-12-22
   EARLIER APPLICATION NUMBER: DE P40 18 988.0
   EARLIER FILING DATE: 1990-06-13
   NUMBER OF SEQ ID NOS: 16
   SOFTWARE: FastSEQ for Windows Version 4.0
      SEQ ID NO 15
      LENGTH: 209
      TYPE: PRT
      ORGANISM: Borrelia burgdorferi
US-09-196-293-15

Query Match   34.7%;   Score 623.5;   DB 4;   Length 209;
Best Local Similarity   69.5%;   Pred. No. 3.5e-42;
Matches 130;   Conservative 24;   Mismatches 32;   Indels 1;   Gaps 1;

```
Qy    1 MACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKA 60
Db   17 ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA 76
```

Best Local Similarity   44.6%;   Pred. No. 6.6e-43;
Matches 166;   Conservative 50;   Mismatches 119;   Indels 37;   Gaps 8;

```
Qy   19 ESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKAIGKKIGNNGLEANQS-KN 77
Db  100 EVFKEDGKTLVSKKVTSKDKSSTEEKFNEKGEVS-EKIITRADGTRLEYTGIKSDSGGKA 158
Qy   78 TSLLSGAYAISDLIAEKLNVLKNE---ELKEKIDTAKQCSTEFTNKLKSEHAVLGLDNLT 134
Db  159 KEVLKGYVLEGTLTAEKTLVVKEGTVTLSKNISKSGEVSVELNDTDSS-AATKKTRAAW 216
Qy  135 DDNAQRAILKKHANKDK---------------GAAELEKLFKAVENLSKAAQDTLKNA 177
Db  217 NDSTSTLTISADSKKTKDLVFLITDGTITVQQYNTAGTSLEGSASEIKNLSE----LKNA 271
Qy  178 VKELTSPIVHGNNSRKDGNASTNSADESVKGPNLTEISKKITESNAVVLAVKEVETLLAS 237
Db  272 LK------MANNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSS 324
Qy  238 IDELATKAIGKKI-GNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAK 296
Db  325 IDEIAAKAIGKKIHQNNGLDTPEYNHNGSLLAGRYAISTLIKQKLDGLIKNEGLKEKIDAAK 384
Qy  297 QCSTEFTNKLKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQ 356
Db  385 KCSETFTNKLKAKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAK 444
Qy  357 DTLKNAVKELTS 368
Db  445 EMLANSVKELTS 456
```

RESULT 6
US-08-158-353-3
; Sequence 3, Application US/08158353
; Patent No. 5620862
; GENERAL INFORMATION:
; APPLICANT: Padula, Steven J.
; TITLE OF INVENTION: Methods for Diagnosing Early Lyme
; TITLE OF INVENTION: Disease
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: MA
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/158,353
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Carroll, Alice O.
; REGISTRATION NUMBER: 33,542
; REFERENCE/DOCKET NUMBER: UCP93-05
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 210 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-158-353-3

Query Match   34.7%;   Score 623.5; DB 1; Length 210;
Best Local Similarity 69.5%;   Pred. No. 3.6e-42;
Matches 130; Conservative 24; Mismatches 32; Indels   1; Gaps   1;

Qy   1 MACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKA   60
         ::||||||||| ||||||||||||||||||||||||||:|::||||||||||||||||| ||
Db   17 ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA   76

Qy   61 IGKKI-GNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTN   119
         |||| | ||||:| |||: ||||||||||||| ||||||| ||||:|: ||
Db   77 IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN   136

Qy   120 KLKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVK   179
         |||:||||| ||||| |||:|| |||:|| | :| ||||||||||:||::| |||:||
Db   137 KLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK   196

Qy   180 ELTSPIV 186
         |||||:|
Db   197 ELTSPVV 203

RESULT 7
US-08-235-836C-30
; Sequence 30, Application US/08235836C
; Patent No. 6248562
; GENERAL INFORMATION:
; APPLICANT: Dunn, John J.
; APPLICANT: Luft, Benjamin J.
; TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising
; TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor
; NUMBER OF SEQUENCES: 144
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brookhaven National Laboratory
; STREET:
; CITY: Upton
; STATE: NY
; COUNTRY: USA
; ZIP: 11973
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/235,836C
; FILING DATE: 29-APR-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/148,191
; FILING DATE: 01-11-93
; ATTORNEY/AGENT INFORMATION:
; NAME: Bogosian, Margaret C.
; REGISTRATION NUMBER: 25,324
; REFERENCE/DOCKET NUMBER: BNL93-28A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 282-7338
; TELEFAX: (516) 282-3729
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 210 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-235-836C-30

Query Match   34.8%;   Score 617.5; DB 4; Length 210;
Best Local Similarity 69.0%;   Pred. No. 1.1e-41;
Matches 129; Conservative 24; Mismatches 33; Indels   1; Gaps   1;

Qy   1 MACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKA   60
         ::||||||||| ||||||||||||||||||||||||||:|::||||||||||||||||| ||
Db   17 ISCNNSGKDGNTSANSADESVGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA   76

Qy   61 IGKKI-GNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTN   119
         |||| | ||||:| |||: ||||||||||||| ||||||| ||||:|: ||
Db   77 IGKKIHQNNGLDTEYNHNGSLLAGRYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN   136

Qy   120 KLKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVK   179
         |||:||||| ||||| |||:|| |||:|| | :| ||||||||||:||::| |||:||
Db   137 KLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK   196

Qy   180 ELTSPIV 186
         |||||:|
Db   197 ELTSPVV 203

RESULT 8
US-08-209-603E-15
; Sequence 15, Application US/0820960E

```
Patent No. 6248538
GENERAL INFORMATION:
   APPLICANT:  FUCHS, RENATE
   APPLICANT:  WILSKE, BETTINA
   APPLICANT:  PREAC-MURSIC, VERA
   APPLICANT:  MOTZ, MANFRED
   APPLICANT:  SOUTSCHECK, ERWIN
   TITLE OF INVENTION:  IMMUNOLOGICALLY ACTIVE PROTEINS
   TITLE OF INVENTION:  FROM BORRELIA BURGDORFERI
   NUMBER OF SEQUENCES:  15
   CORRESPONDENCE ADDRESS:
     ADDRESSEE:  BROOKS HAIDT HAFNER & DELAHUNTY
     STREET:  99 PARK AVENUE
     CITY:  NEW YORK
     STATE:  NY
     COUNTRY:  USA
     ZIP:  10016
   COMPUTER READABLE FORM:
     MEDIUM TYPE:  3.5" FLOPPY DISC
     COMPUTER:  AT&T - IBM COMPATIBLE
     OPERATING SYSTEM:  MS-DOS Version 6.2
     SOFTWARE:  ASCII
   CURRENT APPLICATION DATA:
     APPLICATION NUMBER:  US/08/209,603E
     FILING DATE:  10-MAR-1994
     CLASSIFICATION:  436
   PRIOR APPLICATION DATA:
     APPLICATION NUMBER:  PCT/EP90/02282
     FILING DATE:  21-DEC-1990
     APPLICATION NUMBER:  US 07/862,535
     FILING DATE:  19-JUN-1992
   ATTORNEY/AGENT INFORMATION:
     NAME:  ROBINSON, WILLIAM R.
     REGISTRATION NUMBER:  27,224
     REFERENCE/DOCKET NUMBER:  LKR-9217-A
   TELECOMMUNICATION INFORMATION:
     TELEPHONE:  (212) 697-3355
     TELEFAX:  (212) 557-5635
   INFORMATION FOR SEQ ID NO:  15:
     SEQUENCE CHARACTERISTICS:
       LENGTH:  210
       TYPE:  AMINO ACID
       TOPOLOGY:  LINEAR
     MOLECULE TYPE:  PROTEIN
     DESCRIPTION:  N/A
     HYPOTHETICAL:  N/A
     ANTI-SENSE:  N/A
     FRAGMENT TYPE:  N/A
     ORIGINAL SOURCE:
       ORGANISM:  B. BURGDORFERI
     IMMEDIATE SOURCE:
       LIBRARY:  DSM 5662
     POSITION IN GENOME:  N/A
     FEATURE:
     IDENTIFICATION METHOD:  amino acid analysis
     PUBLICATION INFORMATION:  N/A
US-08-209-603E-15
```

```
Query Match            34.3%;  Score 616.5;  DB 4;  Length 210;
Best Local Similarity  69.0%;  Pred. No. 1.3e-41;
Matches 129;  Conservative 24;  Mismatches 33;  Indels 1;  Gaps 1;

Qy   1  MACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVEVLLASIDELATKA  60
        :|||||||||| |: || ||||  |||||||||||| |||||| : ||||||||||  |||
Db  17  ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVLEVEALLSSIDEIAAKA  76

Qy  61  IGKKI-GNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTN  119
        |||||  ||||||||:|||||||||:|||||||||| || ||||||||||| ||:|| |||
Db  77  IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN  136

Qy 120  KLKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAAELEKLFKAVENLSKAAQDTLKNAVK  179
        | | :: | |||| | | ||| | | ||| :|| |||||||||| |||||:| | |: ||
```

```
Db  137  KLKKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK  196
Qy  180  ELTSPIV  186
         ||||||:|
Db  197  ELTSPVV  203
```

```
RESULT  9
US-08-158-353-4
; Sequence 4, Application US/08158353
; Patent No. 5620862
; GENERAL INFORMATION:
;   APPLICANT:  Padula, Steven J.
;   TITLE OF INVENTION:  Methods for Diagnosing Early Lyme
;   TITLE OF INVENTION:  Disease
;   NUMBER OF SEQUENCES:  7
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE:  Hamilton, Brook, Smith & Reynolds, P.C.
;     STREET:  Two Militia Drive
;     CITY:  Lexington
;     STATE:  MA
;     COUNTRY:  USA
;     ZIP:  02173
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE:  Floppy disk
;     COMPUTER:  IBM PC compatible
;     OPERATING SYSTEM:  PC-DOS/MS-DOS
;     SOFTWARE:  PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER:  US/08/158,353
;     FILING DATE:
;     CLASSIFICATION:  435
;   ATTORNEY/AGENT INFORMATION:
;     NAME:  Carroll, Alice O.
;     REGISTRATION NUMBER:  33,542
;     REFERENCE/DOCKET NUMBER:  UCT93-05
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE:  617-861-6240
;     TELEFAX:  617-861-9540
;   INFORMATION FOR SEQ ID NO:  4:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  212 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-158-353-4
```

```
Query Match            32.7%;  Score 588.5;  DB 1;  Length 212;
Best Local Similarity  68.8%;  Pred. No. 2.1e-39;
Matches 130;  Conservative 20;  Mismatches 36;  Indels 3;  Gaps 3;

Qy   1  MACNNSGKDG-NASANSADESVKGPNLTEISKKITESNAVVLAVKEVEVLLASIDELATK  59
        ::|||||||| :|| ||||||||||||||||||||: ||| ||| || :|||||||||||
Db  17  ISCNNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVEVTLVLSIDELAKK  76

Qy  60  AIGKKI-GNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKN-EELKEKIDTAKQCSTEF  117
        ||||||  |||  |  ||| | || ||||||| |||  | |: ||||| ||| || |||||
Db  77  AIGQKIDNNNGLAALNNQNGSLLAGAYAISTLIITEKLSKLNLEELKTEIAKAKKCSEEF  136

Qy 118  TNKLKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAAELEKLFKAVENLSKAAQDTLKNA  177
        |||| | | | ||  | | ||| | | |||| :|||  ||||||| |||||| |  | |||
Db  137  TNKLKSGHADLGKQDATDHAKAAILKTHATTDKGAKEFKDLFESVEGLLKAAQVALTNS  196

Qy 178  VKELTSPIV  186
        ||||||:|:|
Db  197  VKELTSPVV  205
```

```
RESULT  10
US-08-031-295-2
; Sequence 2, Application US/08031295
```

Patent No. 5530103
GENERAL INFORMATION:
    APPLICANT: LIVEY, Ian
    APPLICANT: DORNER, Freidrich
    TITLE OF INVENTION: METHOD AND COMPOSITION FOR THE
    TITLE OF INVENTION: PREVENTION OF LYME DISEASE
    NUMBER OF SEQUENCES: 3
    CORRESPONDENCE ADDRESS:
        ADDRESSEE: Foley & Lardner
        STREET: 3000 K Street, N.W., Suite 500
        CITY: Washington, D.C.
        COUNTRY: USA
        ZIP: 20007-5109
    COMPUTER READABLE FORM:
        MEDIUM TYPE: Floppy disk
        COMPUTER: IBM PC compatible
        OPERATING SYSTEM: PC-DOS/MS-DOS
        SOFTWARE: Patentin Release #1.0, Version #1.25
    CURRENT APPLICATION DATA:
        APPLICATION NUMBER: US/08/031,295
        FILING DATE: 19930312
        CLASSIFICATION: 435
    PRIOR APPLICATION DATA:
        APPLICATION NUMBER: US 07/903,580
        FILING DATE: 25-JUN-1992
    PRIOR APPLICATION DATA:
        APPLICATION NUMBER: US 07/824,161
        FILING DATE: 22-JAN-1992
    PRIOR APPLICATION DATA:
        APPLICATION NUMBER: US 07/727,245
        FILING DATE: 11-JUL-1991
    ATTORNEY/AGENT INFORMATION:
        NAME: BENT, Stephen A.
        REGISTRATION NUMBER: 29,768
        REFERENCE/DOCKET NUMBER: 30472/142 IMMU
    TELECOMMUNICATION INFORMATION:
        TELEPHONE: (202)672-5300
        TELEFAX: (202)672-5399
        TELEX: 904136
    INFORMATION FOR SEQ ID NO: 2:
        SEQUENCE CHARACTERISTICS:
            LENGTH: 212 amino acids
            TYPE: AMINO ACID
            TOPOLOGY: linear
        MOLECULE TYPE: protein
US-08-031-295-2

Query Match   32.5%; Score 584.5; DB 1; Length 212;
Best Local Similarity   67.2%; Pred. No. 4.3e-39;
tches 127; Conservative 23; Mismatches 36; Indels 3; Gaps 3;

```
Qy   1 MACNNSGKDG-NASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATK 59
        ::|||||| |:||:| ||||||||||||||:|:|||||||||||:|||||||||
Db  17 ISCNNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETLVSSIDELATK 76

Qy  60 AIGKKI-GNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNE-ELKEKIDTAKQCSTEF 117
        ||||| ||||||||||||||||||||| || |||||:|| ||:||||
Db  77 AIGKKIQONNGLGANADKNGSLLAGAYAISTLITEKLKALKNSGELKAKIEDAKKCSEDF 136

Qy 118 TNKLKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNA 177
        | || |||:|| |:|| ||:::: ||||| | ||| ||| |||:|:||:|||:|
Db 137 TKKLAAGHAQLGIDGATDNDSKEAILKTNGTKTKGAEELVKLSESVASLSKAAQEASANS 196

Qy 178 VKELTSPIV 186
        |||||||:|
Db 197 VKELTSPVV 205
```

RESULT 11
US-07-903-580-2
Sequence 2, Application US/07903580
Patent No. 6221363

GENERAL INFORMATION:
    APPLICANT: LIVEY, Ian
    APPLICANT: DORNER, Freidrich
    TITLE OF INVENTION: METHOD AND COMPOSITION FOR THE
    TITLE OF INVENTION: PREVENTION OF LYME DISEASE
    NUMBER OF SEQUENCES: 3
    CORRESPONDENCE ADDRESS:
        ADDRESSEE: Foley & Lardner
        STREET: 1800 Diagonal Road, Suite 500
        CITY: Alexandria
        STATE: Virginia
        COUNTRY: USA
        ZIP: 22313-0299
    COMPUTER READABLE FORM:
        MEDIUM TYPE: Floppy disk
        COMPUTER: IBM PC compatible
        OPERATING SYSTEM: PC-DOS/MS-DOS
        SOFTWARE: Patentin Release #1.0, Version #1.25
    CURRENT APPLICATION DATA:
        APPLICATION NUMBER: US/07/903,580
        FILING DATE: 19920625
        CLASSIFICATION: 435
    PRIOR APPLICATION DATA:
        APPLICATION NUMBER: US 07/824,161
        FILING DATE: 22-JAN-1992
    PRIOR APPLICATION DATA:
        APPLICATION NUMBER: US 07/727,245
        FILING DATE: 11-JUL-1991
    ATTORNEY/AGENT INFORMATION:
        NAME: BENT, Stephen A.
        REGISTRATION NUMBER: 29,768
        REFERENCE/DOCKET NUMBER: 30472/131 IMMU
    TELECOMMUNICATION INFORMATION:
        TELEPHONE: (703)836-9300
        TELEFAX: (703)683-4109
        TELEX: 899149
    INFORMATION FOR SEQ ID NO: 2:
        SEQUENCE CHARACTERISTICS:
            LENGTH: 212 amino acids
            TYPE: AMINO ACID
            TOPOLOGY: linear
        MOLECULE TYPE: protein
US-07-903-580-2

Query Match   32.5%; Score 584.5; DB 4; Length 212;
Best Local Similarity   67.2%; Pred. No. 4.3e-39;
Matches 127; Conservative 23; Mismatches 36; Indels 3; Gaps 3;

```
Qy   1 MACNNSGKDG-NASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATK 59
        ::|||||| |:||:| ||||||||||||||:|:|||||||||||:|||||||||
Db  17 ISCNNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETLVSSIDELATK 76

Qy  60 AIGKKI-GNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNE-ELKEKIDTAKQCSTEF 117
        ||||| ||||||||||||||||||||| || |||||:|| ||:||||
Db  77 AIGKKIQONNGLGANADKNGSLLAGAYAISTLITEKLKALKNSGELKAKIEDAKKCSEDF 136

Qy 118 TNKLKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNA 177
        | || |||:|| |:|| ||:::: ||||| | ||| ||| |||:|:||:|||:|
Db 137 TKKLAAGHAQLGIDGATDNDSKEAILKTNGTKTKGAEELVKLSESVASLSKAAQEASANS 196

Qy 178 VKELTSPIV 186
        |||||||:|
Db 197 VKELTSPVV 205
```

RESULT 12
US-08-235-836C-34
Sequence 34, Application US/08235836C
Patent No. 6248562
GENERAL INFORMATION:
    APPLICANT: Dunn, John J.
    APPLICANT: Luft, Benjamin J.

RESULT 13
US-09-196-293-11
Sequence 11, Application US/09196293
Patent No. 6183755
GENERAL INFORMATION:
APPLICANT: Fuchs, Renate
APPLICANT: Motz, Manfred
APPLICANT: Soutscheck, Erwin
APPLICANT: Wilske, Bettina
APPLICANT: Preac-Mursic, Vera
TITLE OF INVENTION: Active proteins from Borrelia
TITLE OF INVENTION: burgdorferi
FILE REFERENCE: 738.001US2
CURRENT APPLICATION NUMBER: US/09/196,293

TITLE OF INVENTION: No. 62485d2el Chimeric Proteins Comprising
TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor
NUMBER OF SEQUENCES: 144
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brookhaven National Laboratory
STREET:
CITY: Upton
STATE: NY
COUNTRY: USA
ZIP: 11973
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/235,836C
FILING DATE: 29-APR-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/148,191
FILING DATE: 01-11-93
ATTORNEY/AGENT INFORMATION:
NAME: Bogosian, Margaret C.
REGISTRATION NUMBER: 25,324
REFERENCE/DOCKET NUMBER: BNL93-28A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 282-7338
TELEFAX: (516) 282-3729
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 212 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-235-836C-34

Query Match 32.5%; Score 583.5; DB 4; Length 212;
Best Local Similarity 68.3%; Pred. No. 5.1e-39;
Matches 129; Conservative 21; Mismatches 36; Indels 3; Gaps 3;

```
Qy    1 MACNNSGKDG-NASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATK  59
          ||:|| |  | |||||:|||||||||||||||:||:|| | :|||||||||||:|:||||||:
Db   17 ISCSNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETLVLSIDELAKK  76

Qy   60 AIGKKI-GNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKN-EELKEKIDTAKQCSTEF 117
          ||::|| |||| | ::| ||||:|||||||| || :| :|| | |||| ||||:||| ||
Db   77 AIGQKIDNNNGLAALNNQNGSLLAGAYAISTLITEKLSKLKNLEELKTEIAKAKKCSEEF 136

Qy  118 TNKLKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNA 177
         |||||| ||:|| ::|:|||||:|||| |:|:  |||| |||||| |||:|||||:| |:
Db  137 TNKLKSGHADLGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGLLKAAQVALTNS 196

Qy  178 VKELTSPIV 186
         ||||||:|:|
Db  197 VKELTSPVV 205
```

CURRENT FILING DATE: 1998-11-19
EARLIER APPLICATION NUMBER: US 08/209,603
EARLIER FILING DATE: 1994-03-10
EARLIER APPLICATION NUMBER: US 07/862,535
EARLIER FILING DATE: 1992-06-19
EARLIER APPLICATION NUMBER: WO PCT/EP90/02282
EARLIER FILING DATE: 1990-12-21
EARLIER APPLICATION NUMBER: DE P39 42 728.5
EARLIER FILING DATE: 1989-12-22
EARLIER APPLICATION NUMBER: DE P40 18 988.0
EARLIER FILING DATE: 1990-06-13
NUMBER OF SEQ ID NOS: 16
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 11
TYPE: PRT
LENGTH: 212
ORGANISM: Borrelia burgdorferi
US-09-196-293-11

Query Match 32.2%; Score 578.5; DB 4; Length 212;
Best Local Similarity 68.3%; Pred. No. 1.3e-38;
Matches 129; Conservative 19; Mismatches 38; Indels 3; Gaps 3;

```
Qy    1 MACNNSGKDG-NASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATK  59
          ::|||||| | |||||:|||||||||||||||:||:|| | :||||||||||:|:||||||:
Db   17 ISCSNSGKVGILTSTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETLVLSIDELAKK  76

Qy   60 AIGKKI-GNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKN-EELKEKIDTAKQCSTEF 117
          ||::|| |||| | ::| ||||:|||||||| || :| :|| | |||| ||||:||| ||
Db   77 AIGQKIDNNNGLAALNNQNGSLLAGAYAISTLITEKLSKLKNLEELKTEIAKAKKCSEEF 136

Qy  118 TNKLKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNA 177
         |||||| ||:|| ::|:|||||:|||| |:|:  |||| |||||| |||:|||||:| |:
Db  137 TNKLKSGHADLGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGLLKAAQVALTNS 196

Qy  178 VKELTSPIV 186
         ||||||:|:|
Db  197 VKELTSPVV 205
```

RESULT 14
US-08-209-603E-11
Sequence 11, Application US/08209603E
Patent No. 6248538
GENERAL INFORMATION:
APPLICANT: FUCHS, RENATE
APPLICANT: WILSKE, BETTINA
APPLICANT: PREAC-MURSIC, VERA
APPLICANT: MOTZ, MANFRED
APPLICANT: SOUTSCHECK, ERWIN
TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PROTEINS
TITLE OF INVENTION: FROM BORRELIA BURGDORFERI
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROOKS HAIDT HAFFNER & DELAHUNTY
STREET: 99 PARK AVENUE
CITY: NEW YORK
STATE: NY
COUNTRY: USA
ZIP: 10016
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" FLOPPY DISC
COMPUTER: AT&T - IBM COMPATIBLE
OPERATING SYSTEM: MS-DOS Version 6.2
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/209,603E
FILING DATE: 10-MAR-1994
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP90/02282
FILING DATE: 21-DEC-1990

APPLICATION NUMBER: US 07/862,535
FILING DATE: 19-JUN-1992
ATTORNEY/AGENT INFORMATION:
NAME: ROBINSON, WILLIAM R.
REGISTRATION NUMBER: 27,224
REFERENCE/DOCKET NUMBER: LKR-9217-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 697-3355
TELEFAX: (212) 557-5635
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 212
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
HYPOTHETICAL: N/A
DESCRIPTION: N/A
ANTI-SENSE: N/A
FRAGMENT TYPE: N/A
ORIGINAL SOURCE:
ORGANISM: B. BURGDORFERI
IMMEDIATE SOURCE:
LIBRARY: DSM 5662
POSITION IN GENOME: N/A
FEATURE:
IDENTIFICATION METHOD: amino acid analysis
PUBLICATION INFORMATION: N/A
ANTI-SENSE: N/A
FRAGMENT TYPE: INTERNAL
ORIGINAL SOURCE:
US-08-209-603E-11

Query Match 32.2%; Score 578.5; DB 4; Length 212;
Best Local Similarity 68.3%; Pred. No. 1.3e-38;
Matches 129; Conservative 19; Mismatches 38; Indels 3; Gaps 3;

```
Qy    1 MACNNSGKDG-NASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATK 59
          :||||| | |  : | | |||||||||||||: |||| |||||||||| |:|||||| |
Db   17 ISCNNSGKVGILTSTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETLVLSIDELAKK 76

Qy   60 AIGKKI-GNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKN-EELKEIDTAKQCSTEF 117
        ||| ||  ||| ||| |:||||||||:|||| ||| |||||:  |||| ||||| |||
Db   77 AIGQKIDNNNGLAALNNQNGSLLAGAVAISTLITERLKSLKLKNLEELKTEIAKAKKCSEEF 136

Qy  118 TNKLKSEHAVLGLDNLLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNA 177
        |||||| || || ||  | | ||||||| |:| |||| ||| : |||| :: || | |
Db  137 TNKLKSGHADLGKQDAIDDHAKAAILKTHATTDKGAKEFKDLFESVEGLLKAAQVALTNS 196

Qy  178 VKELTSPIV 186
        |||||| :
Db  197 VKELTSPVV 205
```

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/235,836C
FILING DATE: 29-APR-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/148,191
FILING DATE: 01-11-93
ATTORNEY/AGENT INFORMATION:
NAME: Bogosian, Margaret C.
REGISTRATION NUMBER: 25,324
REFERENCE/DOCKET NUMBER: BNL93-28A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 282-7338
TELEFAX: (516) 282-3729
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 207 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-235-8360-36

Query Match 31.3%; Score 563; DB 4; Length 207;
Best Local Similarity 62.6%; Pred. No. 2e-37;
Matches 117; Conservative 34; Mismatches 32; Indels 4; Gaps 3;

```
Qy    1 MACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKA 60
          :||||| | :::: : || ||||||| |||||||: ||| |||||| ||| |:|||| | ||
Db   17 ISCNNSG--GDSASTNPDESAKGPNLTVISKKITDSNAFLLAVKEVEALLSSIDEL-SKA 73

Qy   61 IGKKIGNNG-LEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTN 119
        |||| ||:| : |:| |:||||||||:||| | ||| |:|||| |||: || || |: ||
Db   74 IGKKIKNDGTLDNEANRNESLLAGAVEISKLITQKLSVLNSEELKKKIKEAKDCSEKFTT 133

Qy  120 KLKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVK 179
        ||| ||| ||:  : ||||:|||| |:| ||||| ||||| | :| |||||| | |||
Db  134 KLKDSHAELGIQSVQDDNAKRAILKIHGTKDKGAKELEELFKSLESLSKAAQAALTNSVK 193

Qy  180 ELTSPIV 186
        ||| :||
Db  194 ELTNPVV 200
```

Search completed: March 18, 2002, 09:55:35
Job time: 348 sec

OM protein - protein search, using sw model

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

Run on: March 18, 2002, 10:08:45 ; Search time 621.2 Seconds
                 (without alignments)
                 164.485 Million cell updates/sec

Title: US-09-596-746A-38
Perfect score: 1797
Sequence: 1 MACNNSGKDGNASANSADES..........ENLSKAAQDTLKNAVKELTS 368

Scoring table: BLOSUM62
                Gapop 10.0 , Gapext 0.5

rched: 3148936 seqs, 277657034 residues

Total number of hits satisfying chosen parameters: 3148936

Post-processing: Minimum Match 0%
                Maximum Match 100%
                Listing first 45 summaries

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Database :
```
     Pending_Patents_AA_Main:*
 1: /cgn2_6/ptodata/2/paa/PCTUS_COMB.pep:*
 2: /cgn2_6/ptodata/2/paa/US06_COMB.pep:*
 3: /cgn2_6/ptodata/2/paa/US07_COMB.pep:*
 4: /cgn2_6/ptodata/2/paa/US080_COMB.pep:*
 5: /cgn2_6/ptodata/2/paa/US081_COMB.pep:*
 6: /cgn2_6/ptodata/2/paa/US082_COMB.pep:*
 7: /cgn2_6/ptodata/2/paa/US083_COMB.pep:*
 8: /cgn2_6/ptodata/2/paa/US084_COMB.pep:*
 9: /cgn2_6/ptodata/2/paa/US085_COMB.pep:*
10: /cgn2_6/ptodata/2/paa/US086_COMB.pep:*
11: /cgn2_6/ptodata/2/paa/US087_COMB.pep:*
12: /cgn2_6/ptodata/2/paa/US088_COMB.pep:*
13: /cgn2_6/ptodata/2/paa/US089_COMB.pep:*
14: /cgn2_6/ptodata/2/paa/US090_COMB.pep:*
15: /cgn2_6/ptodata/2/paa/US091_COMB.pep:*
16: /cgn2_6/ptodata/2/paa/US092_COMB.pep:*
17: /cgn2_6/ptodata/2/paa/US093_COMB.pep:*
18: /cgn2_6/ptodata/2/paa/US094_COMB.pep:*
19: /cgn2_6/ptodata/2/paa/US095_COMB.pep:*
20: /cgn2_6/ptodata/2/paa/US096_COMB.pep:*
21: /cgn2_6/ptodata/2/paa/US097_COMB.pep:*
22: /cgn2_6/ptodata/2/paa/US098_COMB.pep:*
23: /cgn2_6/ptodata/2/paa/US099_COMB.pep:*
24: /cgn2_6/ptodata/2/paa/US60_COMB.pep:*
```

SUMMARIES

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1797	100.0	368	19	US-09-596-746A-38	Sequence 38, Appl
2	1797	100.0	368	19	US-09-596-746A-38	Sequence 38, Appl
3	1789	99.6	391	19	US-09-596-746A-76	Sequence 76, Appl
4	1789	99.6	392	19	US-09-596-746A-76	Sequence 76, Appl
5	1538	85.6	368	19	US-09-596-746A-42	Sequence 42, Appl
6	1533	85.3	367	19	US-09-596-746-42	Sequence 42, Appl
7	1530	85.1	391	19	US-09-596-746-74	Sequence 74, Appl
8	1530	85.1	392	19	US-09-596-746A-74	Sequence 74, Appl
9	1490.5	82.9	369	19	US-09-596-746A-40	Sequence 40, Appl

ALIGNMENTS

```
10 1485.5  82.7 368 19 US-09-596-746-40    Sequence 40, Appl
11 1484    82.6 370 19 US-09-596-746A-32   Sequence 32, Appl
12 1482.5  82.5 392 19 US-09-596-746A-72   Sequence 72, Appl
13 1482.5  82.5 391 19 US-09-596-746A-72   Sequence 72, Appl
14 1479    82.3 369 19 US-09-596-746A-32   Sequence 32, Appl
15 1476    82.1 393 19 US-09-596-746-68    Sequence 68, Appl
16 1476    82.1 394 19 US-09-596-746A-68   Sequence 68, Appl
17 1229    68.4 374 19 US-09-596-746-30    Sequence 30, Appl
18 1225.5  68.2 377 19 US-09-596-746A-26   Sequence 26, Appl
19 1225.5  68.2 377 19 US-09-596-746A-30   Sequence 30, Appl
20 1225    68.2 370 19 US-09-596-746A-36   Sequence 36, Appl
21 1224    68.1 373 19 US-09-596-746-26    Sequence 26, Appl
22 1221    67.9 397 19 US-09-596-746-62    Sequence 62, Appl
23 1221    67.9 398 19 US-09-596-746A-62   Sequence 62, Appl
24 1220    67.9 369 19 US-09-596-746A-36   Sequence 36, Appl
25 1219.5  67.9 400 19 US-09-596-746-82    Sequence 82, Appl
26 1219.5  67.9 401 19 US-09-596-746A-82   Sequence 82, Appl
27 1217    67.7 393 19 US-09-596-746-66    Sequence 66, Appl
28 1217    67.7 394 19 US-09-596-746A-66   Sequence 66, Appl
29 1181.5  65.7 375 19 US-09-596-746A-24   Sequence 24, Appl
30 1178    65.6 378 19 US-09-596-746-28    Sequence 28, Appl
31 1178    65.6 371 19 US-09-596-746-24    Sequence 24, Appl
32 1177.5  65.5 374 19 US-09-596-746A-28   Sequence 28, Appl
33 1176.5  65.5 398 19 US-09-596-746-34    Sequence 34, Appl
34 1173.5  65.3 399 19 US-09-596-746A-34   Sequence 34, Appl
35 1173.5  65.3 399 19 US-09-596-746A-60   Sequence 60, Appl
36 1172.5  65.2 370 19 US-09-596-746A-80   Sequence 80, Appl
37 1172    65.2 401 19 US-09-596-746-80    Sequence 80, Appl
38 1172    65.2 402 19 US-09-596-746-64    Sequence 64, Appl
39 1169.5  65.1 394 19 US-09-596-746A-64   Sequence 64, Appl
40 1169.5  65.1 395 19 US-09-596-746A-64   Sequence 64, Appl
41 1167.5  65.0 560 19 US-09-596-746-52    Sequence 52, Appl
42 1167.5  65.0 560 19 US-09-596-746A-52   Sequence 52, Appl
43 1162.5  64.7 385 19 US-09-596-746-56    Sequence 56, Appl
44 1157.5  64.4 384 19 US-09-596-746-56    Sequence 56, Appl
45 1154.5  64.2 408 19 US-09-596-746-78    Sequence 78, Appl
```

RESULT 1
US-09-596-746-38
Sequence 38, Application US/09596746
GENERAL INFORMATION:
 ; APPLICANT: Dattwyler, Raymond J.
 ; APPLICANT: Seinost, Gerald
 ; APPLICANT: Dykhuizen, Daniel
 ; APPLICANT: Luft, Benjamin J.
 ; APPLICANT: Maria J.C. Gomes-Solecki
 ; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
 ; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
 ; FILE REFERENCE: 2631.1002-001
 ; CURRENT APPLICATION NUMBER: US/09/596,746
 ; CURRENT FILING DATE: 2000-06-16
 ; PRIOR APPLICATION NUMBER: US 60/140,042
 ; PRIOR FILING DATE: 1999-06-18
 ; NUMBER OF SEQ ID NOS: 84
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO 38
 ; LENGTH: 368
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: OspC Chimera
US-09-596-746-38

Query Match   100.0%;   Score 1797;   DB 19;   Length 368;
Best Local Similarity   100.0%;   Pred. No. 1.4e-124;
Matches 368;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   1 MACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKA 60

RESULT 2
US-09-596-746A-38
; Sequence 38, Application US/0959746A
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Selnost, Gerald
; APPLICANT: Dykhuizen, Danial
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US-09/596,746A
; CURRENT FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 38
; LENGTH: 368
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: OspC Chimera
US-09-596-746A-38

Query Match   100.0%;   Score 1797;   DB 19;   Length 368;
Best Local Similarity   100.0%;   Pred. No. 1.4e-124;
Matches   368;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

```
Qy    1 MACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKA 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKA 60

Qy   61 IGKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNK 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 IGKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNK 120

Qy  121 LKSEHAVLGLDNLTDDNAQRAILKKHANDKGAAELEKLFKAVENLSKAAQDTLKNAVKE 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 LKSEHAVLGLDNLTDDNAQRAILKKHANDKGAAELEKLFKAVENLSKAAQDTLKNAVKE 180

Qy  181 LTSPIVHGNNSRKDGNASTNSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDE 240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 LTSPIVHGNNSRKDGNASTNSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDE 240

Qy  241 LATKAIGKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCST 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 LATKAIGKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCST 300

Qy  301 EFTNKLKSEHAVLGLDNLTDDNAQRAILKKHANDKGAAELEKLFKAVENLSKAAQDTLK 360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 EFTNKLKSEHAVLGLDNLTDDNAQRAILKKHANDKGAAELEKLFKAVENLSKAAQDTLK 360

Qy  361 NAVKELTS 368
        ||||||||
Db  361 NAVKELTS 368
```

RESULT 3
US-09-596-746-76
; Sequence 76, Application US/0959746
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Selnost, Gerald
; APPLICANT: Dykhuizen, Danial
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US-09/596,746
; CURRENT FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 76
; LENGTH: 391
; TYPE: PRT
; ORGANISM: OspC Chimera
US-09-596-746-76

Query Match   99.6%;   Score 1789;   DB 19;   Length 391;
Best Local Similarity   99.7%;   Pred. No. 6.1e-124;
Matches   366;   Conservative   1;   Mismatches   0;   Indels   0;   Gaps   0;

```
Qy    2 ACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKAI 61
        :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   25 SCNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKAI 84

Qy   62 GKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNKL 121
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   85 GKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNKL 144

Qy  122 KSEHAVLGLDNLTDDNAQRAILKKHANDKGAAELEKLFKAVENLSKAAQDTLKNAVKEL 181
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  145 KSEHAVLGLDNLTDDNAQRAILKKHANDKGAAELEKLFKAVENLSKAAQDTLKNAVKEL 204

Qy  182 TSPIVHGNNSRKDGNASTNSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDEL 241
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  205 TSPIVHGNNSRKDGNASTNSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDEL 264

Qy  242 ATKAIGKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTE 301
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  265 ATKAIGKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTE 324

Qy  302 FTNKLKSEHAVLGLDNLTDDNAQRAILKKHANDKGAAELEKLFKAVENLSKAAQDTLKN 361
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  325 FTNKLKSEHAVLGLDNLTDDNAQRAILKKHANDKGAAELEKLFKAVENLSKAAQDTLKN 384

Qy  362 AVKELTS 368
        |||||||
Db  385 AVKELTS 391
```

RESULT 4
US-09-596-746A-76
; Sequence 76, Application US/0959674A
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seinost, Gerald
; APPLICANT: Dykhuizen, Danial
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746A
; CURRENT FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 76
; LENGTH: 392
; TYPE: PRT
; ORGANISM: ospc Chimera
US-09-596-746A-76

Query Match   99.6%; Score 1789; DB 19; Length 392;
Best Local Similarity   99.7%; Pred. No. 6.1e-124;
Matches 366; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy    2 ACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKAI  61
        :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   26 SCNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKAI  85

Qy   62 GKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNKL 121
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   86 GKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNKL 145

Qy  122 KSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKEL 181
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  146 KSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKEL 205

Qy  182 TSPIVHGNNSRKDGNASTNSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDEL 241
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  206 TSPIVHGNNSRKDGNASTNSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDEL 265

Qy  242 ATKAIGKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTE 301
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  266 ATKAIGKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTE 325

Qy  302 FTNKLKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKN 361
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  326 FTNKLKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKN 385

Qy  362 AVKELTS 368
        |||||||
Db  386 AVKELTS 392
```

RESULT 5
US-09-596-746A-42
; Sequence 42, Application US/0959674A
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seinost, Gerald
; APPLICANT: Dykhuizen, Danial
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746A
; CURRENT FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 42
; LENGTH: 368
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: OspC Chimera
US-09-596-746A-42

Query Match   85.6%; Score 1538; DB 19; Length 368;
Best Local Similarity   86.7%; Pred. No. 2.1e-105;
Matches 319; Conservative 20; Mismatches 27; Indels 2; Gaps 2;

```
Qy    1 MACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKA  60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKA  60

Qy   61 IGKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNK 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 IGKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNK 120

Qy  121 LKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKE 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 LKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKE 180

Qy  181 LTSPIVHGNNSRKDGNASTNSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDE 240
        |||||||||:|||  |||: ||||||||||||||||||||||||||||||||||||||||
Db  181 LTSPIVHGNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDE 240

Qy  241 LATKAIGKKI-GNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKN-EELKEKIDTAKQC 298
        |||:|||| |  |  ||| ::||    ||  | |::|:  |:| ||  |:||||||:|| :
Db  241 LATKAIGKKIQGNGGLAVEAGHNGTLLAGAVTISKLITQKLDGLKNSEKLKEKIENAKKC 300

Qy  299 STEFTNKLKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDT 358
        : :||||||||||||||||||||||||||||||||||||||||||||||||||::|||||
Db  301 SEDFTKLEGEHAQLGTENVTDENAKKAILITDAAAKDKGAAELEKLFKAVENLAKAAKEM 360

Qy  359 LKNAVKEL 366
        | :|||||
Db  361 LANSVKEL 368
```

RESULT 6
US-09-596-746A-42
; Sequence 42, Application US/0959674
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seinost, Gerald
; APPLICANT: Dykhuizen, Danial
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746
; CURRENT FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 42
; LENGTH: 367
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: OspC Chimera
US-09-596-746A-42

Query Match   85.3%; Score 1533; DB 19; Length 367;

RESULT 7
US-09-596-746-74
Sequence 74, Application US/0959674746
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seinost, Gerald
; APPLICANT: Dykhuizen, Danial
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746
; CURRENT FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 74
; LENGTH: 391
; TYPE: PRT
; ORGANISM: ospC Chimera
US-09-596-746-74

Query Match   85.1%; Score 1530; DB 19; Length 391;
Best Local Similarity   86.4%; Pred. No. 8.9e-105;
Matches 317; Conservative   21; Mismatches 27; Indels   2; Gaps   2;

Qy   2 ACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKAI 61
              :!!!
Db   25 SCNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKAI 84

Qy   62 GKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNKL 121
              !!!
Db   85 GKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNKL 144

Qy   122 KSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKEL 181
              !!!
Db   145 KSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKEL 204

Qy   182 TSPIVHGNNSRKDGNASTNSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDEL 241
              !!!!!!!!!!!! !! ! |||| |!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
Db   181 TSPIVHGNNSRKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDEL 240

Qy   242 ATKAIGKKI-GNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKN-EELKEKIDTAKQCS 299
              !!:!|||!!| | ! |:||!:!!|!!!!!!!!!! !:!:!!!|!:| !:!
Db   241 ATKAIGKKIQQNGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKEKIENAKKCS 300

Qy   300 TEFTNKLKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTL 359
              |||!!!!:!!: |
Db   301 EDFTKKLEGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLAKAAKEML 360

Qy   360 KNAVKEL 366
              :!:!!!
Db   361 ANSVKEL 367

RESULT 8
US-09-596-746A-74
Sequence 74, Application US/0959674746A
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seinost, Gerald
; APPLICANT: Dykhuizen, Danial
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746A
; CURRENT FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 74
; LENGTH: 392
; TYPE: PRT
; ORGANISM: ospC Chimera
US-09-596-746A-74

Query Match   85.1%; Score 1530; DB 19; Length 392;
Best Local Similarity   86.4%; Pred. No. 8.9e-105;
Matches 317; Conservative   21; Mismatches 27; Indels   2; Gaps   2;

Qy   2 ACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKAI 61
              :!!!
Db   26 SCNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKAI 85

Qy   62 GKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNKL 121
              !!!
Db   86 GKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNKL 145

Qy   122 KSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKEL 181
              !!!
Db   146 KSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKEL 205

Qy   182 TSPIVHGNNSRKDGNASTNSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDEL 241
              !!!!!!!!!!!! !! ! |||| |!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
Db   206 TSPIVHGNNSRKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDEL 265

Qy   242 ATKAIGKKI-GNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKN-EELKEKIDTAKQCS 299
              !!:!|||!!| | ! |:||!:!!|!!!!!!!!!! !:!:!!!|!:| !:!
Db   266 ATKAIGKKIQQNGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKEKIENAKKCS 325

Qy   300 TEFTNKLKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTL 359
              |||!!!!:!!: |
Db   326 EDFTKKLEGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLAKAAKEML 385

Qy   360 KNAVKEL 366
              :!:!!!
Db   386 ANSVKEL 392

RESULT 9
US-09-596-746A-40
; Sequence 40, Application US/0959746A
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seinost, Gerald
; APPLICANT: Dykhuizen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746A
; CURRENT FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 40
; LENGTH: 369
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: OspC Chimera
US-09-596-746A-40

Query Match   82.9%; Score 1490.5; DB 19; Length 369;
Best Local Similarity 84.3%; Pred. No. 6.9e-102;
Matches 312; Conservative 24; Mismatches 31; Indels 3; Gaps 3;

Qy   1 MACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKA   60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKA   60

Qy   61 IGKKIGNGNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNK   120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 IGKKIGNGNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNK   120

Qy 121 LKSEHAVLGLDNLITDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKE   180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 LKSEHAVLGLDNLITDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKE   180

Qy 181 LTSPIVHGNNSRKDGNASTNSANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDE   240
         |||||||||||||:|||||:|:|||||||||||||||||||||||||||||||||:||||
Db 181 LTSPIVHGNNSKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLTSIDE   240

Qy 241 LATKAIGKKIGNN-GLEANQSKNTSLLSGAYAISDLIAEKLNVLKNE-ELKEKIDTAKQC   298
         || ||||||:|| :||| ::||||||||:|||| ||::||||||| ||:|:|| ::|:
Db 241 LA-KAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAEIEKAKKC   299

Qy 299 STEFTNKLKSEHAVLGLDNLITDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDT   358
         |||| ||||| || | |:|:|||||| |||| |::||||||:|||||:::|||||||::
Db 300 SEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLSKAAKEM   359

Qy 359 LKNAVKELTS   368
         | ::|||||
Db 360 LTNSVKELTS   369

RESULT 10
US-09-596-746-40
; Sequence 40, Application US/0959746
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seinost, Gerald
; APPLICANT: Dykhuizen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746
; CURRENT FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 40
; LENGTH: 368
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: OspC Chimera
US-09-596-746-40

Query Match   82.7%; Score 1485.5; DB 19; Length 368;
Best Local Similarity 84.3%; Pred. No. 1.6e-101;
Matches 311; Conservative 24; Mismatches 31; Indels 3; Gaps 3;

Qy   2 ACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKAI   61
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKAI   60

Qy   62 GKKIGNGNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNKL   121
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 GKKIGNGNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNKL   120

Qy 122 KSEHAVLGLDNLITDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKEL   181
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 KSEHAVLGLDNLITDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKEL   180

Qy 182 TSPIVHGNNSRKDGNASTNSANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLSIDEL   241
         ||||||||||:|||||:|:|||||||||||||||||||||||||||||||||||:||||
Db 181 TSPIVHGNNSKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLTSIDEL   240

Qy 242 ATKAIGKKIGNN-GLEANQSKNTSLLSGAYAISDLIAEKLNVLKNE-ELKEKIDTAKQCS   299
         | ||||||:|| :||| ::||||||||:|||| ||::||||||| ||:|:|| ::|:|
Db 241 A-KAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAEIEKAKKCS   299

Qy 300 TEFTNKLKSEHAVLGLDNLITDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTL   359
         || || ||| || | |:|:|||||| |||| |::||||||:|||||:::|||||||::|
Db 300 EEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLSKAAKEML   359

Qy 360 KNAVKELTS   368
         ::|||||
Db 360 TNSVKELTS   368

RESULT 11
US-09-596-746A-32
; Sequence 32, Application US/0959746A
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seinost, Gerald
; APPLICANT: Dykhuizen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746A
; CURRENT FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 32
; LENGTH: 370
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: OspC Chimera

US-09-596-746A-32

Query Match   82.6%; Score 1484; DB 19; Length 370;
Best Local Similarity   83.8%; Pred. No. 2.1e-101;
Matches   311; Conservative   25; Mismatches   31; Indels   4; Gaps   4;

```
Qy    1 MACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKA   60
        ||||||||||||:||||||||||||||||||||:||||:||||||||:||:||||||.||
Db    1 MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDELA-KA   59

Qy   61 IGKKIGNNG-LEANQSKNTSLLSGAYAISDLIAEKLNVLKNEE-LKEKIDTAKQCSTEFT  118
        |||||:|:| :::||:.:.||||.||.||.||:.:|...|.. :|||||.|.||||.||:
Db   60 IGKKIKNDGSLDNEANRNESLLAGAYTISTLITQKLSKINGSEGLKEKIAAAKCSEEFS   119

Qy  119 NKLKSEHAVIGLDNLTDDNAQRAILKKH-ANKDKGAAELEKLFKAVENLSKAAQDTLKNA  177
        .:||.:|| :|.:| ||:||.:|||| .|..|||.::|||||.|.||.:|||.|::|.|:
Db  120 TKLKDNHAQLGIQGYTDENAKKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANS  179

Qy  178 VKELTSPIVHGHGNNSRKDGNASTNSADESVKGPNLTEISKKITESNAVVLAVKEVETLLAS 237
        |||||||:|||.||||||||||||||||||||||||||||||||||||||||||||||||
Db  180 VKELTSPVVHGNNNSRKDGNASTNSADESVKGPNLTEISKKITESNAVVLAVKEVETLLAS  239

Qy  238 IDELATKAIGKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQ  297
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  240 IDELATKAIGKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQ  299

Qy  298 CSTEFTNKLKSEHAVIGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQD  357
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  300 CSTEFTNKLKSEHAVIGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQD  359

Qy  358 TLKNAVKELTS  368
        |||||||||||
Db  360 TLKNAVKELTS  370
```

RESULT 12
US-09-596-746-72
Sequence 72, Application US/0959674G
GENERAL INFORMATION:
APPLICANT: Dattwyler, Raymond J.
APPLICANT: Seinost, Gerald
APPLICANT: Dykhuizen, Daniel
APPLICANT: Luft, Benjamin J.
APPLICANT: Maria J.C. Gomes-Solecki
TITLE OF INVENTION: Groups of Borrelia burgdorferi and
TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
FILE REFERENCE: 2631.1002-001
CURRENT APPLICATION NUMBER: US/09,596,746
CURRENT FILING DATE: 2000-06-16
PRIOR APPLICATION NUMBER: US 60/140,042
PRIOR FILING DATE: 1999-06-18
SOFTWARE: FastSEQ for Windows Version 4.0
NUMBER OF SEQ ID NOS: 84
SEQ ID NO 72
LENGTH: 392
TYPE: PRT
ORGANISM: ospC Chimera
US-09-596-746-72

Query Match   82.5%; Score 1482.5; DB 19; Length 392;
Best Local Similarity   84.0%; Pred. No. 2.9e-101;
Matches   310; Conservative   25; Mismatches   31; Indels   3; Gaps   3;

```
Qy    2 ACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKAI   61
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   25 SCNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKAI   84

Qy   62 GKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNKL  121
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   85 GKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNKL  144

Qy  122 KSEHAVIGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKEL  181
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  145 KSEHAVIGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKEL  204

Qy  182 TSPIVHGNNSRKDGNASTNSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDEL  241
        ||||||||||.|||||.||||||||||||||||||||||||||||||||||||||.||||
Db  205 TSPIVHGNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLTSIDEL  264

Qy  242 ATKAIGKKIGNN-GLEANQSKNTSLLSGAYAISDLIAEKLNVLKNE-ELKEKIDTAKQCS  299
        |.||||||| .: .:|: .|..|:|||||.|||.|:.:...:| . ||.|||:.|.|||
Db  265 A-KAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAEIEKAKKCS  323

Qy  300 TEFTNKLKSEHAVIGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTL  359
        .:||.:||.||:|.|.|.:|||||.:|||| .:|.||.:|||||||:|.||||||||.||
Db  324 EEFTAKLKGEHTDLGKEGYTDDNAKKAILKTNNDKTKGADELEKLFESVKNLSKAAKEML  383

Qy  360 KNAVKELTS  368
        :|.||||||
Db  384 TNSVKELTS  392
```

RESULT 13
US-09-596-746A-72
Sequence 72, Application US/0959674GA
GENERAL INFORMATION:
APPLICANT: Dattwyler, Raymond J.
APPLICANT: Seinost, Gerald
APPLICANT: Dykhuizen, Daniel
APPLICANT: Luft, Benjamin J.
APPLICANT: Maria J.C. Gomes-Solecki
TITLE OF INVENTION: Groups of Borrelia burgdorferi and
TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
FILE REFERENCE: 2631.1002-001
CURRENT APPLICATION NUMBER: US/09,596,746A
CURRENT FILING DATE: 2000-06-19
PRIOR APPLICATION NUMBER: US 60/140,042
PRIOR FILING DATE: 1999-06-18
SOFTWARE: FastSEQ for Windows Version 4.0
NUMBER OF SEQ ID NOS: 84
SEQ ID NO 72
LENGTH: 393
TYPE: PRT
ORGANISM: ospC Chimera
US-09-596-746A-72

Query Match   82.5%; Score 1482.5; DB 19; Length 393;
Best Local Similarity   84.0%; Pred. No. 2.9e-101;
Matches   310; Conservative   25; Mismatches   31; Indels   3; Gaps   3;

```
Qy    2 ACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKAI   61
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   26 SCNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKAI   85

Qy   62 GKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNKL  121
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   86 GKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNKL  145

Qy  122 KSEHAVIGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKEL  181
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  146 KSEHAVIGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKEL  205

Qy  182 TSPIVHGNNSRKDGNASTNSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDEL  241
        ||||||||||.|||||.||||||||||||||||||||||||||||||||||||||.||||
Db  206 TSPIVHGNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLTSIDEL  265

Qy  242 ATKAIGKKIGNN-GLEANQSKNTSLLSGAYAISDLIAEKLNVLKNE-ELKEKIDTAKQCS  299
        |.||||||| .: .:|: .|..|:|||||.|||.|:.:...:| . ||.|||:.|.|||
Db  266 A-KAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAEIEKAKKCS  324

Qy  300 TEFTNKLKSEHAVIGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTL  359
        .:||.:||.||:|.|.|.:|||||.:|||| .:|.||.:|||||||:|.||||||||.||
Db  325 EEFTAKLKGEHTDLGKEGYTDDNAKKAILKTNNDKTKGADELEKLFESVKNLSKAAKEML  384

Qy  360 KNAVKELTS  368
        :|.||||||
Db  385 TNSVKELTS  392
```

```
RESULT 14
US-09-596-746-32
Sequence 32, Application US/09596746
GENERAL INFORMATION:
  APPLICANT: Dattwyler, Raymond J.
  APPLICANT: Luft, Benjamin J.
  APPLICANT: Dykhuizen, Danial
  APPLICANT: Maria J.C. Gomes-Solecki
  TITLE OF INVENTION: Groups of Borrelia burgdorferi and
  TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
  FILE REFERENCE: 2631.1002-001
  CURRENT APPLICATION NUMBER: US/09/596,746
  CURRENT FILING DATE: 2000-06-16
  PRIOR APPLICATION NUMBER: US 60/140,042
  PRIOR FILING DATE: 1999-06-18
  NUMBER OF SEQ ID NOS: 84
  SOFTWARE: FastSEQ for Windows Version 4.0
  SEQ ID NO 32
  LENGTH: 369
  TYPE: PRT
  ORGANISM: Artificial Sequence
  FEATURE:
  OTHER INFORMATION: OspC Chimera
US-09-596-746-32

Query Match        82.3%;  Score 1479;  DB 19;  Length 369;
Best Local Similarity   83.8%;  Pred. No. 4.9e-101;
Matches 310;  Conservative 25;  Mismatches 31;  Indels  4;  Gaps 4;

Qy     2 ACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKAI 61
           || ||||||| ||||||||||||:|||||||||||||||||| :||||||||| :|||||
Db     1 ACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDELA-KAI 59

Qy    62 GKKIGNNG-LEANQSKNTSLLSGAYAISDLIAEKLNVLKNEE-LKEKIDTAKQCSTEFTN 119
           |||| || |||||| |||||||||||||||||||| :|:|| |||||||||| |
Db    60 GKKIKNDGSLDNEANRNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKCSEEFST 119

Qy   120 KLKSEHAVLGLDNLTDDNAQRAILKKH-ANKDKGAAELEKLFKAVENLSKAAQDTLKNAV 178
           ||| | ::|| :|||:||:| ::::|: |:||:::::|||||||||||||||||:|
Db   120 KLKDNHAQLGIQGVTDENAKKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANSV 179

Qy   179 KELTSPIVHGNNSRKDGNASTNSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASI 238
           |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db   180 KELTSPVVHGNNSRKDGNASTNSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASI 239

Qy   239 DELATKAIGKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQC 298
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   240 DELATKAIGKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQC 299

Qy   299 STEFTNKLKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDT 358
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   300 STEFTNKLKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDT 359

Qy   359 LKNAVKELTS 368
           |||||||
Db   360 LKNAVKELTS 369
```

```
RESULT 15
US-09-596-746-68
Sequence 68, Application US/09596746
GENERAL INFORMATION:
  APPLICANT: Dattwyler, Raymond J.
  APPLICANT: Selnost, Gerald
  APPLICANT: Dykhuizen, Danial
```

```
  APPLICANT: Luft, Benjamin J.
  APPLICANT: Maria J.C. Gomes-Solecki
  TITLE OF INVENTION: Groups of Borrelia burgdorferi and
  TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
  FILE REFERENCE: 2631.1002-001
  CURRENT APPLICATION NUMBER: US/09/596,746
  CURRENT FILING DATE: 2000-06-16
  PRIOR APPLICATION NUMBER: US 60/140,042
  PRIOR FILING DATE: 1999-06-18
  NUMBER OF SEQ ID NOS: 84
  SOFTWARE: FastSEQ for Windows Version 4.0
  SEQ ID NO 68
  LENGTH: 393
  TYPE: PRT
  ORGANISM: OspC Chimera
US-09-596-746-68

Query Match        82.1%;  Score 1476;  DB 19;  Length 393;
Best Local Similarity   83.5%;  Pred. No. 8.9e-101;
Matches 309;  Conservative 26;  Mismatches 31;  Indels  4;  Gaps 4;

Qy     2 ACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKAI 61
           || ||||||| ||||||||||||:|||||||||||||||||| :||||||||| :|||||
Db    25 SCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDELA-KAI 83

Qy    62 GKKIGNNG-LEANQSKNTSLLSGAYAISDLIAEKLNVLKNEE-LKEKIDTAKQCSTEFTN 119
           |||| || |||||| |||||||||||||||||||| :|:|| |||||||||| |
Db    84 GKKIKNDGSLDNEANRNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKCSEEFST 143

Qy   120 KLKSEHAVLGLDNLTDDNAQRAILKKH-ANKDKGAAELEKLFKAVENLSKAAQDTLKNAV 178
           ||| | ::|| :|||:||:| ::::|: |:||:::::|||||||||||||||||:|
Db   144 KLKDNHAQLGIQGVTDENAKKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANSV 203

Qy   179 KELTSPIVHGNNSRKDGNASTNSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASI 238
           |||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   204 KELTSPVVHGNNSRKDGNASTNSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASI 263

Qy   239 DELATKAIGKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQC 298
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   264 DELATKAIGKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQC 323

Qy   299 STEFTNKLKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDT 358
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   324 STEFTNKLKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDT 383

Qy   359 LKNAVKELTS 368
           |||||||||
Db   384 LKNAVKELTS 393

Search completed: March 18, 2002, 10:08:45
Job time: 973 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw.model

Run on: March 18, 2002, 09:58:23 ; Search time 81.02 seconds
                                          (without alignments)
                                          290.158 Million cell updates/sec

Title: US-09-596-746a-38
Perfect score: 1797
Sequence: 1 MACNNSGKDGNASANSADES.........ENLSKAAQDTLKNAVKELTS 368

Scoring table: BLOSUM62
               Gapop 10.0 , Gapext 0.5

Searched: 303921 seqs, 63882009 residues

Total number of hits satisfying chosen parameters: 303921

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   Pending_Patents_AA_New:*
               1: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pep:*
               2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep:*
               3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep:*
               4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep:*
               5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep:*
               6: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep1:*
               7: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep:*
               8: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	588.5	32.7	212	6	US-09-974-992-7	Sequence 7, Appli
2	583	32.4	211	6	US-09-974-992-5	Sequence 5, Appli
3	567	31.6	207	6	US-09-974-992-3	Sequence 3, Appli
4	191	10.6	837	6	US-09-815-242-5883	Sequence 5883, Ap
5	191	10.6	837	7	US-10-072-851-5883	Sequence 5883, Ap
6	191	10.6	875	6	US-09-815-242-13080	Sequence 13080, A
7	191	10.6	875	7	US-10-072-851-13080	Sequence 13080, A
8	182	10.1	2434	6	US-09-815-242-5835	Sequence 5835, Ap
9	182	10.1	2434	7	US-10-072-851-5835	Sequence 5835, Ap
10	182	10.1	6281	6	US-09-815-242-12996	Sequence 12996, A
11	182	10.1	6281	7	US-10-072-851-12996	Sequence 12996, A
12	173.5	9.7	1881	6	US-09-815-242-7646	Sequence 7646, Ap
13	173.5	9.7	1881	7	US-10-072-851-7646	Sequence 7646, A
14	171.5	9.5	2076	6	US-09-815-242-15590	Sequence 15590, A
15	171.5	9.5	2076	7	US-10-072-851-15590	Sequence 15590, A
16	171.5	9.5	2186	6	US-09-815-242-5815	Sequence 5815, Ap
17	171.5	9.5	2186	7	US-10-072-851-5815	Sequence 5815, A
18	167	9.3	1144	6	US-09-708-427-15046	Sequence 15046, A
19	167	9.3	1304	6	US-09-708-427-15045	Sequence 15045, A
20	167	9.3	1313	6	US-09-708-427-15044	Sequence 15044, A
21	163.5	9.1	1561	6	US-09-708-427-13229	Sequence 13229, A
22	163.5	9.1	1686	6	US-09-708-427-13228	Sequence 13228, A
23	163.5	9.1	1730	6	US-09-708-427-13227	Sequence 13227, A
24	162.5	9.0	1014	6	US-09-708-427-19883	Sequence 19883, A
25	162.5	9.0	1018	6	US-09-708-427-19882	Sequence 19882, A
26	162.5	9.0	1269	6	US-09-708-427-19881	Sequence 19881, A
27	162.5	9.0	2056	6	US-09-614-150-4824	Sequence 4824, Ap
28	159	8.8	2437	6	US-09-815-242-5834	Sequence 5834, Ap
29	159	8.8	2437	7	US-10-072-851-5834	Sequence 5834, Ap
30	158	8.8	5795	6	US-09-815-242-12610	Sequence 12610, A
31	158	8.8	5795	7	US-10-072-851-12610	Sequence 12610, A
32	157.5	8.8	2086	6	US-09-815-242-5639	Sequence 5639, Ap
33	157.5	8.8	2086	7	US-10-072-851-5639	Sequence 5639, Ap
34	157	8.7	1639	6	US-09-614-150-6213	Sequence 6213, Ap
35	156.5	8.7	1013	6	US-09-614-150-9159	Sequence 9159, Ap
36	152.5	8.5	1128	6	US-10-032-585-7627	Sequence 7627, A
37	152.5	8.5	1128	7	US-10-072-851-15571	Sequence 15571, A
38	152	8.5	2025	6	US-09-815-242-5703	Sequence 5703, Ap
39	152	8.5	2025	7	US-10-072-851-5703	Sequence 5703, Ap
40	152	8.5	3158	6	US-09-815-242-12611	Sequence 12611, A
41	152	8.5	3158	7	US-10-072-851-12611	Sequence 12611, A
42	151.5	8.4	2013	7	US-09-614-150-13758	Sequence 13758, A
43	149.5	8.3	1765	7	US-10-037-182-8	Sequence 8, Appli
44	149.5	8.3	1786	5	US-09-936-275-6	Sequence 6, Appli
45	149.5	8.3	1786	6	US-09-873-676-113	Sequence 113, App

ALIGNMENTS

RESULT 1
US-09-974-992-7

Sequence 7, Application US/09974992

GENERAL INFORMATION:

; APPLICANT: Mathiesen, Marianne J.
; APPLICANT: Theisen, Michael
; APPLICANT: Holm, Arne
; APPLICANT: Ostergaard, Soren
; TITLE OF INVENTION: Novel OspC-derived peptide fragments
; FILE REFERENCE: 459-666P
; CURRENT APPLICATION NUMBER: US/09/974,992
; CURRENT FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US/09/180,089
; PRIOR FILING DATE: 1999-05-13
; PRIOR APPLICATION NUMBER: PCT/DK97/00203
; PRIOR FILING DATE: 1997-05-02
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 7
;   LENGTH: 212
;   TYPE: PRT
;   ORGANISM: Borrelia afzelii
US-09-974-992-7

Query Match   32.7%;   Score 588.5; DB 6; Length 212;
Best Local Similarity 68.8%;   Pred. No. 2.9e-32;
Matches 130; Conservative 20; Mismatches 36; Indels 3; Gaps 3;

Qy   1 MACNNSGKDG-NASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATK 59
            ::|||||||| ||||| |||:|||||||||||||::|| |||||||||||||:|||||||
Db   17 ISCNNSGKGGDSASTNPADESARGPNLTEISKKITDSNAFVLAVKEVETLVLSIDELAKK 76

Qy   60 AIGKKI-GNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKN-EELKEKIDTAKQCSTEF 117
            |||||| || |||:||||| ||:||:|||||||||||||||| :|||||| || ||| :|
Db   77 AIGQKIDNNNGLAALNNQNGSLLAGAYAISTLITEKLSKLKNLEELKTEIAKAKKCSEEF 136

Qy   118 TNKLKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNA 177
            |||||| |: |||||||||||||||||||||:|| || ||||||||||:|||||||||:
Db   137 TNKLKSGHADLGKDOATDDHAKAAILKTHATTDKGAKEFKDLFESVEGLLKAAQVALTNS 196

Qy   178 VKELTSPIV 186
            ||||||| :
Db   197 VKELTSPVV 205

RESULT 2
US-09-974-992-5

Sequence 5, Application US/09974992
; GENERAL INFORMATION:
; APPLICANT: Mathiesen, Marianne J.
; APPLICANT: Theisen, Michael
; APPLICANT: Holm, Arne
; APPLICANT: Ostergaard, Soren
; TITLE OF INVENTION: Novel OspC-derived peptide fragments
; FILE REFERENCE: 459-666P
; CURRENT APPLICATION NUMBER: US/09/974,992
; CURRENT FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: 09/180,089
; PRIOR FILING DATE: 1999-05-13
; PRIOR APPLICATION NUMBER: PCT/DK97/00203
; PRIOR FILING DATE: 1997-05-02
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 211
; TYPE: PRT
; ORGANISM: Borrelia burgdorferi
US-09-974-992-5

Query Match   32.4%; Score 583; DB 6; Length 211;
Best Local Similarity   67.5%; Pred. No. 6.8e-32;
Matches 129; Conservative 24; Mismatches 30; Indels 8; Gaps 5;

```
Qy   1 MACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKA  60
         ::|||||| |  |:::: |||||:||||||||||||||||||||||||| ||:|||||:||
Db  17 ISCNNSG---GDSASTNPDESAKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDELA-KA  75

Qy  61 IGKKIGNNGL---EANQSKNTSLLSGAYAISDLIAEKLNVLKNEE-LKEKIDTAKQCSTE  116
         :|||||||||   || |:||||||||||||||||:|||||||:| ||||||| |||||||
Db  76 IGKKIKNDGSLGDEANH--NESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEE  133

Qy 117 FTNKLKSEHAVLGILDNLTDDNAQRAILKKH-ANKDKGAAELEKLFKAVENLSKAAQDTLK 175
         |:||||:|||:||||:|||:||||:||||  |:|||||:||||||::||:||||||:| |
Db 134 FSTKLKDNHAQLGIGQVTDENAKKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLA  193

Qy 176 NAVKELTSPIV 186
         |:||||||||:|
Db 194 NSVKELTSPVV 204
```

RESULT 3
US-09-974-992-3
; Sequence 3, Application US/09974992
; GENERAL INFORMATION:
; APPLICANT: Mathiesen, Marianne J.
; APPLICANT: Theisen, Michael
; APPLICANT: Holm, Arne
; APPLICANT: Ostergaard, Soren
; TITLE OF INVENTION: Novel OspC-derived peptide fragments
; FILE REFERENCE: 459-666P
; CURRENT APPLICATION NUMBER: US/09/974,992
; CURRENT FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: 09/180,089
; PRIOR FILING DATE: 1999-05-13
; PRIOR APPLICATION NUMBER: PCT/DK97/00203
; PRIOR FILING DATE: 1997-05-02
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 207
; TYPE: PRT
; ORGANISM: Borrelia garinii
US-09-974-992-3

Query Match   31.6%; Score 567; DB 6; Length 207;
Best Local Similarity   63.1%; Pred. No. 7.6e-31;
Matches 118; Conservative 33; Mismatches 32; Indels 4; Gaps 3;

```
Qy   1 MACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKA  60
         :|||||| |::|||||:  |||||:||||||||||:||| ::|||||||| ||:|||||:
Db  17 ISCNNSG---GDSASTNPDESAKGPNLTVISKKITDSNAFLLAVKEVEALLSSIDEL-SKA  73

Qy  61 IGKKIGNNG--LEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTN 119
         ||||| | |  |:||  |:|||:|||| ||.|| ||.||| ::||||||  |||||:| ||
Db  74 IGKKIKNDGTLDNEARNESLIAGAYEISKLIITQKLSVLNSEELKEKIKPAKDCSEKFTT  133

Qy 120 KLKSEHAVLGILDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVK 179
         |||:.||:|||::|.  ||||:||||:| .|||||:||| ||| :||||||||| ::|:||
Db 134 KLKDSHAELGIGQSVQDDNAKKAILKTHGTKDGAKELEELFKSLESLSKAAQAALTNSVK  193

Qy 180 ELTSPIV 186
         |||.|:|
Db 194 ELTNPVV 200
```

RESULT 4
US-09-815-242-5883
; Sequence 5883, Application US/09815242
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US/09/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5883
; LENGTH: 837
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-5883

Query Match   10.6%; Score 191; DB 6; Length 837;
Best Local Similarity   22.4%; Pred. No. 3.8e-05;
Matches 91; Conservative 84; Mismatches 148; Indels 84; Gaps 19;

```
Qy   4 NNSGKDGNASANSADE---SVKGPNLTEISKKITESNAV--VLAVKE-VETLLASI----  53
         |:   :.:| :|:::    ||||||| |  .  :|.:||   .:.:: |.|.|:|
Db 100 NHNLEVAKQNANTAIDGLISLNGPQKAKLKEQVGQATTLPNVQTVRDNAQTLNTAMKGLR  159

Qy  54 DELATKAIGKKIGNNGLEANQSKNTSLLSGAYAISDLIAE---KLNVLKNEELKEKIDT  109
         |.:|.:|  .|:  |  .:|.: : :..|.:|.   ::|:   ||||:|:||::|.:||
Db 160 DSIANEAT-IKAQQNTDASQNKQTDYNSAVTAAKAILGQTTSPSMNAQEINQAKDQV-T  217

Qy 110 AKQCSTEFTNKLK-----SEHAVLGILDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVE 164
         |||    .||.|      ||||:||.|.|||  |||:.:.:|  .|||:||| ||||||.
Db 218 AKQQALNGQENLRTAQTNAKQHLNGLSDLTD--AQKDAVKRQI---EGATHVNEVTQAQN  272
```

RESULT 5
10-072-851-5883
Sequence 5883, Application US/10072851
GENERAL INFORMATION:
; APPLICANT: Carr, Grant J.
; APPLICANT: Xu, H. Howard
; APPLICANT: Foulkes, J. Gordon
; APPLICANT: Zamudio, Carlos
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Roemer, Terry
; APPLICANT: Jiang, Bo
; APPLICANT: Boone, Charles
; APPLICANT: Bussey, Howard
; TITLE OF INVENTION: Methods for Identifying the Target of a Compound which Inhibits C
; TITLE OF INVENTION: Proliferation
; FILE REFERENCE: ELITRA.028A
; CURRENT APPLICATION NUMBER: US/10,072,851
; CURRENT FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 15811
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5883
; LENGTH: 837
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-072-851-5883

Query Match   10.6%;   Score 191; DB 7; Length 837;
Best Local Similarity 22.4%;   Pred. No. 3.8e-05;
Matches   91; Conservative 84; Mismatches 148; Indels 84; Gaps 19;

```
Qy    4 NNSGKDGNASANSADE---SVKGPNLTEISKKITESNAV--VLAVKE-VETLLASI----  53
Db  100 NHNLEVAQKONANTAIDGLTSLNGPQKAKLKEQVGQOATTLPNVQTVRDNAQTLNTAMKGLR 159
Qy   54 DELATKAIGKKIGNNGLEANQSKNTSLLSGAYAISDLIAE---KLNVLKNEELKEKIDT  109
Db  160 DSIANEAT-IKAGONYTDASONKQTDYNSAVTAAKAIIGQTTSPSMNAQEINQAKDQV-T  217
Qy  110 AKQCSTEFTNKLK----SEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVE  164
Db  218 AKQQALNGQENLRTAQTNAKQHLNGLSDLTD--AQKDAVKRQI--EGATHVNEVTQAQN  272
Qy  165 NLS--KAAQDTLKNAVKELTSPIVHGNNSRKDGNASTNSADESVKGPNLTEISKKITESN 222
Db  273 NADALNTAMTNLKNGIQD------QNTIKQG-VNFTDADEA------KRN           309
Qy  223 AVVLAVKEVETLL-------ASIDELATKAIGKKIGNNGLEANQSKNTSLLSGAYAISDL 275
Db  310 AYTNAVTQAEQILNKAQGPNTSKDGVETALENVQRAKNELNGNQNVANAKTTAKNALNNL 369
Qy  276 IAEKLNVLKNEELKEKID------------TAKQCSTEFTNLKLSEHAVLGLDN-LTDD  321
Db  370 TS--INNAQKEALKSQIEGATTVAGVNQVSTTASELNTAMSN------------LQNGINDE 417
Qy  322 NAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKELTS            368
Db  418 AATKAALNGTQNLEKAKQHANTAIDGLSHLTNAQKEALKQLVQQSTT            464
```

RESULT 6
US-09-815-242-13080
Sequence 13080, Application US/09815242
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13080
; LENGTH: 875
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-13080

Query Match   10.6%;   Score 191; DB 6; Length 875;
Best Local Similarity 22.4%;   Pred. No. 4e-05;
Matches   91; Conservative 84; Mismatches 148; Indels 84; Gaps 19;

```
Qy    4 NNSGKDGNASANSADE---SVKGPNLTEISKKITESNAV--VLAVKE-VETLLASI----  53
Db  137 NHNLEVAKQNANTAIDGLTSLNGPQKAKLKEQVGQOATTLPNVQTVRDNAQTLNTAMKGLR 196
Qy   54 DELATKAIGKKIGNNGLEANQSKNTSLLSGAYAISDLIAE---KLNVLKNEELKEKIDT  109
Db  197 DSIANEAT-IKAGONYTDASONKQTDYNSAVTAAKAIIGQTTSPSMNAQEINQAKDQV-T  254
Qy  110 AKQCSTEFTNKLK----SEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVE  164
Db  255 AKQQALNGQENLRTAQTNAKQHLNGLSDLTD--AQKDAVKRQI--EGATHVNEVTQAQN  309
Qy  165 NLS--KAAQDTLKNAVKELTSPIVHGNNSRKDGNASTNSADESVKGPNLTEISKKITESN 222
Db  310 NADALNTAMTNLKNGIQD------QNTIKQG-VNFTDADEA------KRN           346
Qy  223 AVVLAVKEVETLL-------ASIDELATKAIGKKIGNNGLEANQSKNTSLLSGAYAISDL 275
```

```
Db   347  AYTNAVTQAEQILNKAQGPNTSKDGVETALENVQRAKNELNGNQNVANAKTTAKNALNNL 406
               :  |||    :|  ||||:|| :||
Qy   276  IAEKLNVLKNEELKEKID------------------TAKQCSTEFTNKLKSEHAVLGLDN-LTDD 321
               ||  ||  |::|:
Db   407  TS--INNAQKEALKSQIEGATTVAGVNQVSTTASELNTAMSN-------------LQNGINDE 454

Qy   322  NAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKELTS 368
               :|:  ||| |: :||
Db   455  AATKAALNGTQNLEKAKQHANTAIDGLSHLTNAQKEALKQLVQQSTT 501
```

RESULT 7
US-10-072-851-13080
; Sequence 13080, Application US/10072851
; GENERAL INFORMATION:
; APPLICANT: Carr, Grant J.
; APPLICANT: Xu, H. Howard
; APPLICANT: Foulkes, J. Gordon
; APPLICANT: Zamudio, Carlos
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Roemer, Terry
; APPLICANT: Jiang, Bo
; APPLICANT: Boone, Charles
; APPLICANT: Bussey, Howard
; TITLE OF INVENTION: Methods for Identifying the Target of a Compound which Inhibits C
; TITLE OF INVENTION: Proliferation
; FILE REFERENCE: ELITRA.028A
; CURRENT APPLICATION NUMBER: US/10/072,851
; CURRENT FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 15811
; SEQ ID NO 13080
; LENGTH: 875
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-072-851-13080

Query Match   10.6%; Score 191; DB 7; Length 875;
Best Local Similarity 22.4%; Pred. No. 4e-05;
Matches   91; Conservative   84; Mismatches 148; Indels   84; Gaps   19;

```
Qy     4  NNSGKDGNASANSADE--------SVKGPNLTEISKKITESNAV--VLAVKE-VETLLASI---- 53
               ||||     :|:  :|| ||    ::| |:|  |:: | ||   ||::  ||:||::|
Db   137  NHNLEVAQONANTAIDGLTFSLNGPQKAKLKEQVGQOATTLPNVQTVRDNAQTLNTAMKGLR 196

Qy    54  DELATKAIGKKIGNNGLEANQSKNTSLLSGAYAISDLIAE-----KLNVLKNEELKEKIDT 109
               |  ||:|:| ||||||:|||||||:||||||||||:| ||    ||||||||||:||| :
Db   197  DSIANEAT-IKAGQNYTDASQNKQTDYNSAVTAAKAIIGQTTSPSMNAQEINQAKDQV-T 254

Qy   110  AKQCSTEFTNKLK-------SEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVE 164
               |||||  | | |        |:|||||||||||:|||:|| |||||||:|||||||||||:
Db   255  AKQQALNGQOENLRTAQTNAKQHLNGLSLSDLTD--AQKDAVKRQI--EGATHVNEVTQAQN 309

Qy   165  NLS--KAAQDTLKNAVKELTSPIVHGNSRKDGNASTNSADESVKGPNLTEISKKITESN 222
               ||:  |||| :||||:|||||| | |||:||||:|:||||||:|| ||:|:|||:||:|
Db   310  NADALNTAMTNLKNGIQD-------QNTIKQG-VNFTDADEA--KRN 346

Qy   223  AVVLAVKEVETLL-----ASIDELATKAIGKKIGNNGLEANQSKNTSLLSGAYAISDL 275
               :  ||:|||| ||     |:|   ||:||:|  ||||||||||||:|||||||||:|
Db   347  AYTNAVTQAEQILNKAQGPNTSKDGVETALENVQRAKNELNGNQNVANAKTTAKNALNNL 406

Qy   276  IAEKLNVLKNEELKEKID------------------TAKQCSTEFTNKLKSEHAVLGLDN-LTDD 321
               |  ||  |::|:|          :|  |:
Db   407  TS--INNAQKEALKSQIEGATTVAGVNQVSTTASELNTAMSN-------------LQNGINDE 454
```

RESULT 8
US-09-815-242-5835
; Sequence 5835, Application US/09815242
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SEQ ID NO 5835
; LENGTH: 2434
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-5835

Query Match   10.1%; Score 182; DB 6; Length 2434;
Best Local Similarity 22.2%; Pred. No. 0.00057;
Matches 107; Conservative   81; Mismatches 161; Indels 132; Gaps   24;

```
Qy     4  NNSGKDGNASANSADE--------SVKGPNLTEISKKITESNAV--VLAVKE-VETLLASI---- 53
               ||||     :|:  :|| ||    ::| |:|  |:: | ||   ||::  ||:||::|
Db   702  NHNLEVAQONANTAIDGLTFSLNGPQKAKLKEQVGQOATTLPNVQTVRDNAQTLNTAMKGLR 761

Qy    54  DELATKAIGKKIGNNGLEANQSKNTSLLSGAYAISDLIAE-----KLNVLKNEELKEKIDT 109
               |  ||:|:| ||||||:|||||||:||||||||||:| ||    ||||||||||:||| :
Db   762  DSIANEAT-IKAGQNYTDASQNKQTDYNSAVTAAKAIIGQTTSPSMNAQEINQAKDQV-T 819

Qy   110  AKQCSTEFTNKLK-------SEHAVLGLDNLTDDNAQRAILKK-----HANK-DKGAAELE 157
               |||||  | | |        |:|||||||||||:|||:|| ||
Db   820  AKQQALNGQOENLRTAQTNAKQHLNGLSLSDLTD--AQKDAVKRQIEGATHVNEVTQAQNNAD 877

Qy   158  KLFKAVENLSKAAQD--TLK--------------------NAVKELTSPI--VHGNSRKD-- 194
               ||  |:  ||||||   :|                     |||||||| |  :| ||| |
Db   878  ALNTAMTNLKNGIQDQONTIKQGVNFTDADEAKRNAYTNAVTQAEQILNKAQGPNTSKDGV 937

Qy   195  ----------GNASTNSADESVKGP--NLTEISKKITES------------NA 223
               ||:|:||||||:|| ||:|:|||:||:      |:
Db   938  ETALENVQRAKNELNGNQNVANAKTTAKNALNNLTSINNAQKEALKSQIEGATTVAGVNQ 997

Qy   224  VVLAVKEVETLLASI-----DELATKAIGKKIGNNGLEANQSKNTSLLSGAYAISDLI-- 276
               ::  | |   |::       ||||||:|:| ||||||:|||||||:|||||||||:|
Db   998  VSTTASELNTAMSNLQNGINDEAATKAAQKY-----TDADREKQTAYNDAVTAAKTLDK 1052
```

RESULT 9
US-10-072-851-5835

Sequence 5835, Application US/10072851

; GENERAL INFORMATION:
; APPLICANT: Carr, Grant J.
; APPLICANT: Xu, H. Howard
; APPLICANT: Foulkes, J. Gordon
; APPLICANT: Zamudio, Carlos
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Roemer, Terry
; APPLICANT: Jiang, Bo
; APPLICANT: Boone, Charles
; APPLICANT: Bussey, Howard
; TITLE OF INVENTION: Methods for Identifying the Target of a Compound which Inhibits O
; TITLE OF INVENTION: Proliferation
; FILE REFERENCE: ELITRA.028A
; CURRENT APPLICATION NUMBER: US/10/072,851
; CURRENT FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 15811
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5835
; LENGTH: 2434
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-072-851-5835

Query Match   10.1%;   Score 182;   DB 7;   Length 2434;
Best Local Similarity   22.2%;   Pred. No. 0.00057;
Matches 107;   Conservative 81;   Mismatches 161;   Indels 132;   Gaps 24;

```
Qy    277 ------------AEKLNVLK-----NEELKEKIDTAKQ------------CSTEFTNKL 306
              :|:|     |:|||
Db   1053 TAGSNDNKAAVEQALQRVNTARTALNGDERLNEAKNTAKQQVATMSHLTDAQKANLTSQI 1112

Qy    307 KSEHAVLGLDNL------TDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKN 361
              ||:|                ||  | :||  | |:|  :| ||
Db   1113 ESGTTVAGVQGIQANAGTLDQAMNQLRQSIASKD-ATKSSEDYQDANADLQNAYNDAVTN 1171

Qy    362 A 362
            |
Db   1172 A 1172
```

RESULT 10
US-09-815-242-12996

Sequence 12996, Application US/09815242

; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12996
; LENGTH: 6281
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-12996

Query Match   10.1%;   Score 182;   DB 6;   Length 6281;
Best Local Similarity   22.2%;   Pred. No. 0.0019;
Matches 107;   Conservative 81;   Mismatches 161;   Indels 132;   Gaps 24;

```
Db    998 VSTPASELNTAMSNLQNGINDEAATKAAQKY-----TDADREKQTAYNDAVTAARTLLDK 1052

Qy    277 ------------AEKLNVLK-----NEELKEKIDTAKQ------------CSTEFTNKL 306
              :|:|     |:|||
Db   1053 TAGSNDNKAAVEQALQRVNTARTALNGDERLNEAKNTAKQQVATMSHLTDAQKANLTSQI 1112

Qy    307 KSEHAVLGLDNL------TDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKN 361
              ||:|                ||  | :||  | |:|  :| ||
Db   1113 ESGTTVAGVQGIQANAGTLDQAMNQLRQSIASKD-ATKSSEDYQDANADLQNAYNDAVTN 1171

Qy    362 A 362
            |
Db   1172 A 1172
```

```
Qy      4 NNSGKDGNASANSADE---SVKGPNLTEISKKITESNAV--VLAVKE-VETLLASI---- 53
            |:|      :|||:|    | |::|||    :|   :||||: :|      |:
Db    702 NHNLEVAKQNANTAIDGLTSLNGPQRAKLKEQVGQOATTLPNVQTVRDNAQTLNTAMKGLR 761

Qy     54 DELATKAIGKKIGNNGLEANQSKNTSLLSGAYAISDLIAE---KLNVLKNEELKEKIDT 109
            |:|:| ||  |||||||||||||:|:|||||||||:|||:    |:|||:|||||||:
Db    762 DSIANEAT-IKAGQNYDASONKQTDYNSAVTAAKAIIGQTTSPSMNAQEINQAKDQV-T 819

Qy    110 AKQCSTEFTNKLK-----SEHAVLGLDNLTDDNAQRAILKK-----HANK-DKGAAELE 157
            |:||       :        ||:|      :| : :|:||       |||  :||||||
Db    820 AKQQALNGQOENLRTAQTNAKQHLNGLSDLTD--AQKDAVKRQIEGATHVNEVTQAQNNAD 877

Qy    158 KLFKAVENLSKAAQD--TLK-----------NAVKELTSPI--VHGNNSRKD-- 194
            || | ||    ||||  |||           |:||:||||   ||||||||
Db    878 ALNTAMTNLKNGIQDQNTIKQGVNFTDADEARKNAYTNAVTQAEQILNKAQGNTSKDGV 937

Qy    195 --------GNASTNSADESVKGP---NLTEISKKITES------ NA 223
                    |:||::|||||:|||    :||||:|||:          ||
Db    938 ETALENVQRAKNELNGNQNVANAKTTAKNALNNLTSINNAQREALKSQIEGATTVAGVNQ 997

Qy    224 VVLAVKEVETLLASI------DELATKAIGKKIGNNGLEANQSKNTSLLSGAYAISDLI-- 276
            |:|: : :    ||||      ||||||||:|||||||||||||||:|:|||||||||
Db    998 ... 1599
```

RESULT 11
US-10-072-851-12996
Sequence 12996, Application US/10072851
; GENERAL INFORMATION:
; APPLICANT: Carr, Grant J.
; APPLICANT: Xu, H. Howard
; APPLICANT: Foulkes, J. Gordon
; APPLICANT: Zamudio, Carlos
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Roemer, Terry
; APPLICANT: Jiang, Bo
; APPLICANT: Boone, Charles
; APPLICANT: Bussey, Howard
; TITLE OF INVENTION: Methods for Identifying the Target of a Compound which Inhibits C
; TITLE OF INVENTION: Proliferation
; FILE REFERENCE: ELITRA.028A
; CURRENT APPLICATION NUMBER: US/10/072,851
; CURRENT FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: 60/257,636
; PRIOR FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 15811
; SQ ID NO 12996
; LENGTH: 6281
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
; SOFTWARE: FastSEQ for Windows version 4.0

Query Match 10.1%;   Score 182;   DB 7;   Length 6281;
Best Local Similarity 22.2%;   Pred. No. 0.0019;
Matches 107;   Conservative 81;   Mismatches 161;   Indels 132;   Gaps 24;

```
Qy     4 NNSGKDGNASANSADE---SVKGPNLTEISKKITESNAV--VLAVKE-VETLLASI----   53
           :||:|            |:||:|||| |::| :|| |   |:||| ||||||||
Db  1481 NHHLEVAKQNANTAIDGLTSLNGFDQAKLKEQVQGOATLPNVQTVRDNAQTLINTAMKGLR 1540

Qy    54 DELATKAIGKKIGNNGLEANQSKNTSLLSGAYAISDLIE------KLNVLKNEELKEKIDT  109
          :||| ||:||  |: |||  ||||||||||||| ||  |     |:|  |:|  |||| |
Db  1541 DSIANEAT-IKAGONYTDASQNKQTDYNSAVTAAKAIGQTTSPSMNAQEINQAKDQV-T   1598

Qy   110 AKQCSTEFTNKLK-----SEHAVLGLDNLITDDNAQRAILKK-----HANK-DKGAAELE   157
          |||  ||||||:|     ||||:|||||| ||  ||||||||     |||| ||| |||
Db  1599 AKQOALNGQENLRFAQTNAKQHLNGLSDLITD--AQKDAVKRQIEGATHVNEVTQAQNNAD 1656

Qy   158 KLFKAVENLSKAAQD--TLK--------------------------NAVKELTSPI--VHGNNSRKD--  194
           |||:|||:|:|:|    |||                          |::||:|:|    ||||||:|
Db  1657 ALNTAMTNLKNGIQDQNTIKQGVNFTDADEAKRNAYTNAVTQAEQILNKAQGPNTSKDGV 1716

Qy   195 -------------GNASTNSADESVKGP--NLTEISKKITES------------NA   223
                       ||:||||:|:||:|    |||||:|||||            |
Db  1717 ETALENVQRAKNELNGNQNVANAKTTAKNALNNLTSINNAQKEALKSQIEGATTVAGVNQ 1776

Qy   224 VVLAVKEVETLLASI--------DELATKAIGKKIGNNGLEANQSKNTSLLSGAYAISDLI-- 276
          |:|:||||:||||||        |||||||||:| |: ||| ||||||||||||||||||
Db  1777 VSTTASELNTAMSNLQNGINDEAATKAAQKY-----TDADREKQTAYNDAVTAAKTLLDK   1831

Qy   277 -------AEKLNVLK-----NEELKEKIDTAKQ-------------CSTEFTNKL   306
                 |||||:||     ||||:|||||||             |||||||:|
Db  1832 TAGSNDNKAAVEQALQRVNTAKTALNGDERLNEAKNTAKQQVATMSHLTDAQKANLITSQI 1891

Qy   307 KSEHAVLGLDNL-----TDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKN   361
          :||||:|||:||     ||||||||||:||:||||||||||||||||||||||||:||:|
Db  1892 ESGTTVAGVQGIQANAGTLDQAMNQLRQSIASKD-ATKSSEDYQDANADLQNAYNDAVTN  1950

Qy   362 A   362
          |
Db  1951 A  1951
```

RESULT 12
US-10-032-585-7646
Sequence 7646, Application US/10032585
; GENERAL INFORMATION:
; APPLICANT: Terry, Roemer D.
; APPLICANT: Bo, Jiang
; APPLICANT: Charles, Boone
; APPLICANT: Howard, Bussey
; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
; CURRENT APPLICATION NUMBER: US/10/032,585
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 8000
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 7646
; LENGTH: 1881
; TYPE: PRT
; ORGANISM: Candida albicans
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (1881)..(1881)
; OTHER INFORMATION: X=any amino acid
US-10-032-585-7646

Query Match 9.7%;   Score 173.5;   DB 7;   Length 1881;
Best Local Similarity 25.6%;   Pred. No. 0.0015;
Matches 112;   Conservative 55;   Mismatches 150;   Indels 121;   Gaps 25;

```
Qy     2 ACNNSGKD-----GNASANSAD--ESVK------------GPNLTEISKKITESNAVVLAVKE  45
          | :||:||     ||:|||:|    |:|            |||||||||||||||:|||||
Db  1470 ALTKSSKDLEVCGNQKSELQDSLKSVKSELKNFENKYNOQETTSLKDEIEEKQ------KE  1523

Qy    46 VETLLASIDELATKAIGKKIGNNGLEANQSKNTSLLSGAYAISDLI--AEKLNVLKNEE   102
          |:||| |||||| ||:|||:| |||:|:|||:||||||||||||||  |:|||:|:
Db  1524 IVTLQTELKDRISEVEKER----AMLSENSETVIKEY--SDRIKSLESKINSIKENH    1574

Qy   103 LKEKIDTAKQCSTEFTNKLKSEHAVLGLDN------LTDDNAQ--RAILKKH-----   146
          ::|:||||:| ||::  :|   |:|  ||:      |||||||  |||||:
Db  1575 SKE-ITTHNEQKT----SLKQDIAKLSQDHESAQTQLEDKENQLKELKASLEKHNTESAT 1629

Qy   147 --ANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKELTSPIVHGNSRKDGNASTNS--   201
            |||||||||||||||:|:|||::|:|||:|  |::|||:||:|||||||||:|
Db  1630 SIEEKNQIKELSETIKSLKTELKTSGDALKQSQKEY------KILKTKNSDTESKLE   1681

Qy   202 ---------------ADESVKGPNLTEISKKITESNAVVLAVKEVETL----LASIDELA 242
                         |||:|||:|||||:||||||:|:||:|:|:||    |||:|:|
Db  1682 KQLEELEKVKSDLQTADEKLKGITERELALK----------SELETVKNSGLSTSELA   1730
```

```
Qy   243  TKAIGKKIGNNGLEANQSKNTSLLSG--AYAISDL-----IAEKLNVLKNEELKEK---  291
Db  1731  --ALTKTV-KSLE-KEKEELQFLSGNKSKELEDVIQKHSDISEKLKAL-TDELKEKTKQ 1784
Qy   292  IDTAKQCSTEFTNKLKSEHAVLGLDNLTDDNAQRAILKK-HANKDKGAAELEKLFKAVEN  350
Db  1785  FDDSKKKLTELENDLTSTK-----KELETEKTQTSKFKNLEERKDKEIVKLNKELELLKN 1839
Qy   351  LSKAAQDTLKNAVKELTS  368
Db  1840  DNSGAKKELSEKVSKLES 1857
```

RESULT 13
US-10-072-851-15590
Sequence 15590, Application US/10072851
GENERAL INFORMATION:
APPLICANT: Carr, Grant J.
APPLICANT: Xu, H. Howard
APPLICANT: Foulkes, J. Gordon
APPLICANT: Zamudio, Carlos
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Roemer, Terry
APPLICANT: Jiang, Bo
APPLICANT: Boone, Charles
APPLICANT: Bussey, Howard
TITLE OF INVENTION: Methods for Identifying the Target of a Compound which inhibits c
TITLE OF INVENTION: Proliferation
FILE REFERENCE: ELITRA.028A
CURRENT APPLICATION NUMBER: US/10,072,851
CURRENT FILING DATE: 2002-02-08
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
NUMBER OF SEQ ID NOS: 15811
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 15590
LENGTH: 1881
TYPE: PRT
ORGANISM: Candida albicans
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (1881)..(1881)
OTHER INFORMATION: X=any amino acid
US-10-072-851-15590

Query Match   9.7%;   Score 173.5; DB 7; Length 1881;
Best Local Similarity 25.6%; Pred. No. 0.0015;
Matches 112; Conservative 55; Mismatches 150; Indels 121; Gaps 25;

```
Qy     2  ACNNSGKD----GNASANSAD--ESVK----------GPNLTEISKKITESNAVLAVKE   45
Db  1470  ALTKSSKDLEVCGNQKSELODSLKSVKSELKNFENKYNQETTSLKDEIEEKQ------KE 1523

Qy    46  VETLLASIDELATKAIGKKIGNNGLEANQSKNTSLLSGAYAISDLI---AEKLNVLKNEE  102
Db  1524  IVTLQTELKDRISEVEKER------AMLSENSETVIKEY--SDKIKSLESKINSIKENH  1574

Qy   103  LKEKIDTAKQCSTEFTNKLKSEHAVLGLDN-----LTTDDNAQ----RAILKKH-----  146
Db  1575  SKE-TTTHNEQKT----SLKODIAKLSQDHESAQTQLEDKENQLKELKASLEKHNTESAT 1629

Qy   147  --ANKDGAAELEKLFKAVENLSKAAQDTLKNAVKELTSPIVHGNNSRKDGNASTNS---  201
Db  1630  SIEERKNNQIKELSETIKSLKTELKTSGDALKQSOKEY------KTLKTNSDTESKLE  1681

Qy   202  ------------ADESVKGPNLTEISKKITESNAVVLAVKEVETL---LASIDELA  242
```

```
Db  1682  KQLEELEVKSDLQTADEKLKGITEREIALK-----------SELETVKNSGLSTTSELA 1730
Qy   243  TKAIGKKIGNNGLEANQSKNTSLLSG-AYAISDL------------IAEKLNVLKNEELKEK---  291
Db  1731  --ALTKTV-KSLE-KEKEELQFLSGNKSKELEDVIQKHSDISEKLKAL-TDELKEKTKQ 1784
Qy   292  IDTAKQCSTEFTNKLKSEHAVLGLDNLTDDNAQRAILKK-HANKDKGAAELEKLFKAVEN  350
Db  1785  FDDSKKKLTELENDLTSTK-----KELETEKTQTSKFKNLEERKDKEIVKLNKELELLKN 1839
Qy   351  LSKAAQDTLKNAVKELTS  368
Db  1840  DNSGAKKELSEKVSKLES 1857
```

RESULT 14
US-09-815-242-5815
Sequence 5815, Application US/09815242
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
TITLE OF INVENTION: Prokaryotes
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 5815
LENGTH: 2076
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-815-242-5815

Query Match   9.5%;   Score 171.5; DB 6; Length 2076;
Best Local Similarity 20.0%; Pred. No. 0.0023;
Matches 85; Conservative 78; Mismatches 174; Indels 87; Gaps 14;

```
Qy    20  SVKGPNLTEISKKITESNAVL-AVKEVETLLASIDELATKA---IGKKIGNNGLEA--   72
Db  1132  TAKKOEINQNTNATTEEKQVALNQVDQELATAININIQADTNAEVDQAQLGTKAINAIQ 1191

Qy    73  ------------NQSKNTSLL---SGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTE  116
Db  1192  PNIVKKPAALAQINQHYNAKLAEINATPDATNDEKNAAINTL-NQDRQQAIESIKQAN- 1248

Qy   117  FTNKLKSEHAVLGLDNLTDDNAQRAILKKHIANKDKGAAELEKLFKAVENLSKAAQDTLKN  176
Db  1249  -TNAEVDQAAATVAENNI--DAVQVDVVKKQAARDKITAEVAKRIEAVKQTPNATDEEKQA 1305

Qy   177  AVKELTS-----PIVHGNNSRKDGNASTN------SADESVKGPNLTEISKKITES  221
```

RESULT 15
10-072-851-5815
Sequence 5815, Application US/10072851
GENERAL INFORMATION:
APPLICANT: Carr, Grant J.
APPLICANT: Xu, H. Howard
APPLICANT: Foulkes, J. Gordon
APPLICANT: Zamudio, Carlos
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Roemer, Terry
APPLICANT: Jiang, Bo
APPLICANT: Boone, Charles
APPLICANT: Bussey, Howard
TITLE OF INVENTION: Methods for Identifying the Target of a Compound which Inhibits C
TITLE OF INVENTION: Proliferation
FILE REFERENCE: ELITRA.028A
CURRENT APPLICATION NUMBER: US/10/072,851
CURRENT FILING DATE: 2002-02-08
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
NUMBER OF SEQ ID NOS: 15811
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 5815
   LENGTH: 2076
   TYPE: PRT
   ORGANISM: Staphylococcus aureus
10-072-851-5815

Query Match   9.5%;   Score 171.5; DB 7;   Length 2076;
Best Local Similarity   20.0%;   Pred. No. 0.0023;
Matches   85;   Conservative   78;   Mismatches   174;   Indels   87;   Gaps   14;

```
Qy   20 SVKGPNLTEISKKITESNAVVL--AVKEVETLLASIDELATKA--IGKKIGNNGLEA--  72
         : ||::  : || ::: |  | :|:|: |::|:: |::| |  |: || :
Db 1132 TAKKQEINQNTNATTEEKQVALNQVDQELATAINNINQADTNAEVDQAQQLGTKAINAIQ 1191

Qy   73 ------------NQSKNTSLL---SGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTE 116
                    :| :  |||| : ::|::|::| |:: |: |:  | |:
Db 1192 PNIVKKPAALAQINQHYNAKLAEINATPDATNDEKNAAINTL-NQDRQAIESIKQAN-- 1248

Qy  117 FTNKLKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKN 176
         | |::|: ::|| |:|:::| |: :|||| ||::|:: | :||
Db 1249 -TNAEVDQAATVAENNI---DAVQDVVKKQAARDKITAEVAKRIEAVKQTPNATDEEKQA 1305

Qy  177 AVKELTS-----PIVHGNNSRKDGNASTN--------SADESVKGPNLTEISKKITES 221
         || :|||       |:|:| :|:  :|           :| ::|||| |:|:::
Db 1306 AVNQINQLKDQAINQINQNQTNDQVDTTTNQAVNAIDNVEAEVVIKPRAIADIEKAVKEK 1365

Qy  222 N-------------------AVVLAVKEVETLLASIDELATKAIGKKIGNNGLEANQSKNTSL 265
         |                    | | :|:|: |::|:: |::| |  |: || :
Db 1366 QQQIDNSLDSTDNEKEVASQALAKEKEKALAAIDQAQTNSQVNQAATNGVSA-----IKI 1420

Qy  266 LSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCS-----------TEFTNKLKSEHAVLG 314
         || : ::|::|::| |:: |: |:  | |:            | |:
Db 1421 IQPETKVKPAAREKIN-QKANELRAKINQDKEATAEERQVALDKINEFVNQAMTDITNNR 1479

Qy  315 LDNLTDDNAQRAI----------LKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVK 364
         |:|:::| |: :          :|||| ||::|:: | :|| || :||
Db 1480 TNQQVDDTTSQALDSIALVTPDHIVRAAARDAVKQQYEAKKREIEQAEHATDEEKQVALN 1539

Qy  365 ELTS 368
         ::|:
Db 1540 QLAN 1543
```

Search completed: March 18, 2002, 09:58:28
Job time: 386 sec

OM protein - protein search, using sw model

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

Run on: March 18, 2002, 09:56:48 ; Search time 68.77 Seconds
                   (without alignments)
                   407.623 Million cell updates/sec

Title: US-09-596-746A-38
Perfect score: 1797
Sequence: 1 MACNNSGKDGNASANSADES............ENLSKAAQDTLKNAVKELTS 368

Scoring table: BLOSUM62
                   Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 200000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database : PIR_68:*
        1: piri:*
        2: pir2:*
        3: pir3:*
        4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	Match DB	ID	Description
1	901	50.1	209	2	I40142	outer surface prot
2	661	36.8	193	2	S70287	outer surface prot
3	646	35.9	211	2	I40145	outer surface prot
4	642	35.7	193	2	S70279	outer surface prot
5	639	35.6	191	2	I40153	outer surface prot
6	623.5	35.5	210	2	G70218	outer surface prot
7	622	34.7	209	2	G70218	outer surface prot
8	619.5	34.5	210	2	S69917	outer surface prot
9	619	34.4	193	2	S69927	outer surface prot
10	614	34.2	191	2	S70286	outer surface prot
11	612.5	34.1	212	2	S70284	outer surface prot
12	605.5	33.7	194	2	S69922	outer surface prot
13	605	33.7	178	2	S70268	outer surface prot
14	605	33.7	178	2	I40104	outer surface prot
15	603.5	33.6	212	2	I40125	outer surface prot
16	596.5	33.2	212	2	I40144	outer surface prot
17	589	32.8	193	2	S70274	outer surface prot
18	588.5	32.7	193	2	S20543	outer surface prot
19	588	32.7	211	2	I40268	outer surface prot
20	587	32.7	211	2	S69918	outer surface prot
21	586	32.6	211	2	S69932	outer surface prot
22	584.5	32.5	212	2	S70254	outer surface prot
23	581	32.3	193	2	S70280	outer surface prot
24	579.5	32.2	212	2	S69921	outer surface prot
25	577	32.1	211	2	I40278	outer surface prot
26	575	32.0	193	2	S70265	outer surface prot
27	572	31.8	193	2	S70276	outer surface prot
28	569.5	31.7	177	2	I40129	outer surface prot
29	569	31.7	207	2	I40271	outer surface prot
30	567	31.6	200	2	I40122	outer surface prot
31	567	31.6	207	2	S69919	outer surface prot
32	567	31.6	207	2	S37727	outer surface prot
33	566.5	31.5	190	2	S70261	outer surface prot
34	566.5	31.5	194	2	S70277	outer surface prot
35	566.5	31.5	212	2	I40279	outer surface prot
36	566	31.5	209	2	I40281	outer surface prot
37	565.5	31.5	194	2	S70289	outer surface prot
38	565	31.4	211	2	I40276	outer surface prot
39	564	31.4	207	2	I40277	outer surface prot
40	562.5	31.3	209	2	S69916	outer surface prot
41	562	31.2	180	2	I40110	outer surface prot
42	560.5	31.2	180	2	S70263	outer surface prot
43	560	31.1	189	2	I40108	outer surface prot
44	559	31.1	203	2	I40124	outer surface prot
45	558	31.1	178	2	I40124	outer surface prot

ALIGNMENTS

RESULT   1
I40142
outer surface protein C precursor - Lyme disease spirochete
outer surface protein C precursor (Lyme disease spirochete)
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 21-Jul-2000
C;Accession: I40142; S70283
A;Status: Preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-209 <RES>
R;Stevenson, B.; Bockenstedt, L.K.; Barthold, S.W.
Infect. Immun. 62, 3568-3571, 1994
A;Title: Expression and gene sequence of outer surface protein C of Borrelia burgdorf
A;Cross-references: EMBL:U04240; NID:g2314877; PIDN:AAC45538.1; PID:g434658
R;Livey, I.; Gibbs, C.P.; Schuster, R.; Dorner, F.
Mol. Microbiol. 18, 257-269, 1995
A;Title: Evidence for lateral transfer and recombination in OspC variation in Lyme di
A;Reference number: S70255; MUID:96296448
A;Accession: S70283
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 19-209 <LIV>
A;Cross-references: EMBL:L42894; NID:g858722; PIDN:AAB37002.1; PID:g1695219
A;Experimental source: strain 28691
C;Genetics:
A;Gene: ospC
C;Superfamily: Lyme disease spirochete surface protein C

Query Match   50.1%;   Score 901;   DB 2;   Length 209;
Best Local Similarity   98.9%;   Pred. NO. 2.2e-35;
Matches 184;   Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy   1 MACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKA 60
            ::||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   17 ISCNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKA 76

Qy   61 IGKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNK 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   77 IGKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNK 136

Qy   121 LKSEHAVLGLDNLTDDNAQRAILKHHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKE 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   137 LKSEHAVLGLDNLTDDNAQRAILKHHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKE 196

Qy   181 LTSPIV 186
            ||||||
Db   197 LTSPIV 202

RESULT   2

S70287
outer surface protein C - Lyme disease spirochete
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 26-May-2000
C;Accession: S70287
R;Livey, I.; Gibbs, C.P.; Schuster, R.; Dorner, F.
Mol. Microbiol. 18, 257-269, 1995
A;Title: Evidence for lateral transfer and recombination in OspC variation in Lyme disea
A;Reference number: S70287
A;Accession: S70287
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-193 <LIV>
A;Cross-references: EMBL:L42895; NID:g858723; PIDN:AAB37003.1; PID:g1695220
A;Experimental source: strain 28354
A;Genetics:
A;Gene: ospC
C;Superfamily: Lyme disease spirochete surface protein C

Query Match   36.8%; Score 661; DB 2; Length 193;
Best Local Similarity   74.7%;
Matches 139; Conservative 20; Mismatches 25; Indels 2; Gaps 2;

Qy   3 CNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKAIG 62
            ||||||||| ||||||||||||||||||||||||:|||||||||||||||||||||||||
Db   1 CNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDELATKAIG 60

Qy   63 KKI-GNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKN-EELKEKIDIAKQCSTEFTNK 120
            ||| ||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db   61 KKIQQONGGLAVEAGHNGTLLLAGAYTISKLITQKLDGLKNSEKLKEKIENAKKCSEDFTKK 120

Qy   121 LKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKE 180
            |:||||||||||||||||||||||||:||||| |||||||||||||||||:||:|||||||
Db   121 LEGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLAKAAAKEMLANSVKE 180

Qy   181 LTSPIV 186
            |||||
Db   181 LTSPIV 186

RESULT 3
I40145
outer surface protein C precursor - Lyme disease spirochete
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 21-Jul-2000
C;Accession: I40145
R;Stevenson, B.; Barthold, S.W.
FEMS Microbiol. Lett. 124, 367-372, 1994
A;Title: Expression and sequence of outer surface protein C among North American isolate
A;Reference number: I40143; MUID:95154673
A;Accession: I40145
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-211 <RES>
A;Cross-references: EMBL:U04282; NID:g2314881; PIDN:AAC45540.1; PID:g4434666
A;Gene: ospC
C;Superfamily: Lyme disease spirochete surface protein C

Query Match   35.9%; Score 646; DB 2; Length 211;
Best Local Similarity   72.3%;
Matches 136; Conservative 18; Mismatches 32; Indels 2; Gaps 2;

Qy   1 MACNNSGKDGN-ASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATK 59
          ||:|||||||| |||||||||||||||||||||||||:|||||||||||||||||||||||
Db   17 ISCNNSGKDGNAASTNPADESVKGPNLTEISKKITDSNAVVLAVKEVGALITSIDELATK 76

Qy   60 AIGKKI-GNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDIAKQCSTEFT 118
          |||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   77 AIGKKIHQONNGLDTENNHNGSLLAGAYAISTLITQKLGGLKNEELKEKIAAVKKCSEEFT 136

Qy   119 NKLKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAV 178
          |||| |||| || :| |||||||||||:||:|||||||||||||||::||||| ||||
Db   137 NKLKSSHTELGKQDADDDAKKAILRTHNTKDKGAAELDKLFKAVENLSKAAKEMLSNSV 196

Qy   179 KELTSPIV 186
          |||||||:|
Db   197 KELTSPVV 204

RESULT 4
S70279
outer surface protein C - Lyme disease spirochete
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 26-May-2000
C;Accession: S70279
R;Livey, I.; Gibbs, C.P.; Schuster, R.; Dorner, F.
Mol. Microbiol. 18, 257-269, 1995
A;Title: Evidence for lateral transfer and recombination in OspC variation in Lyme di
A;Reference number: S70255; MUID:96296448
A;Accession: S70279
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-193 <LIV>
A;Cross-references: EMBL:L42898; NID:g858729; PIDN:AAB37007.1; PID:g1695223
A;Experimental source: strain 25015
A;Genetics:
A;Gene: ospC
C;Superfamily: Lyme disease spirochete surface protein C

Query Match   35.7%; Score 642; DB 2; Length 193;
Best Local Similarity   72.6%;
Matches 135; Conservative 16; Mismatches 33; Indels 2; Gaps 2;

Qy   3 CNNSGKDGN-ASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKAI 61
            ||||||||| ||||||||||||||||||||||||:||||||||||||||||||||||||
Db   1 CNNSGKDGNAASTNPADESVKGPNLTEISKKITDSNAVVLAVKEVEALLTSIDELATKAI 60

Qy   62 GKKI-GNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDIAKQCSTEFTNK 120
          |||| ||||||||||||||||||||||||||||||:|||||||||||||||||||||||
Db   61 GKKIHQONNGLDTENNHNGSLLAGAYAISTLITQKLGGLKNEELKEKIAAVKKCSEEFTNK 120

Qy   121 LKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKE 180
          |:|| |||| || :| |||||||||||||||||||||||:|||||||:||||| |||||
Db   121 LKSSHTELGKQDADDDAKKAILRTHNTKDKGAAELDKLFKPVENLSKAAKEMLSNSVKE 180

Qy   181 LTSPIV 186
            |||:|
Db   181 LTSPVV 186

RESULT 5
I40153
outer surface protein C - Lyme disease spirochete (fragment)
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 26-May-2000
C;Accession: I40153
R;Fung, B.P.; McHugh, G.L.; Leong, J.M.; Steere, A.C.
Infect. Immun. 62, 3213-3221, 1994
A;Title: Humoral immune response to outer surface protein C of Borrelia burgdorferi i
A;Reference number: I40153; MUID:94314437
A;Accession: I40153
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-191 <RES>
A;Cross-references: EMBL:U08284; NID:g469561; PIDN:AAA21460.1; PID:g469562
A;Genetics:
A;Gene: ospC
C;Superfamily: Lyme disease spirochete surface protein C

Query Match   35.6%; Score 639; DB 2; Length 191;
Best Local Similarity   73.9%;

RESULT 6

```
G70218
...er surface protein C - Lyme disease spirochete
...pecies: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 21-Jul-2000
C;Accession: G70218; I40269; S37726; S70281
R;Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White
  son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt,
  ; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
  Nature 390, 580-586, 1997
A;Authors: Smith, H.O.; Venter, J.C.
A;Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.
A;Reference number: A70100; MUID:98065943
A;Accession: G70218
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-210 <KLE>
A;Cross-references: GB:AE000792; NID:g3253098; PIDN:AAC66329.1; PID:g2689901; TIGR:BBB19
J. Fukunaga, M.; Hamase, A.
J. Clin. Microbiol 33, 2415-2420, 1995
A;Title: Outer surface protein C gene sequence analysis of Borrelia burgdorferi sensu la
A;Reference number: I40269; MUID:96025162
A;Accession: I40269
A;Molecule type: DNA
A;Residues: 1-210 <RES>
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Cross-references: GB:D49497; NID:g7070092; PIDN:BAA08457.1; PID:g9769684
R;Jauris-Helpke, S.; Fuchs, R.; Motz, M.; Preac-Mursic, V.; Schwab, E.; Soutschek, E.; W
  Med. Microbiol. Immunol. 182, 37-50, 1993
A;Title: Genetic heterogenity of the genes coding for the outer surface protein C (OspC)
A;Reference number: S37726; MUID:93268136
A;Accession: S37726; MUID:93268136
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-210 <JAU>
A;Cross-references: EMBL:X69596; NID:g311391; PIDN:CAA49306.1; PID:g311392
R;Livey, I.; Gibbs, C.P.; Schuster, R.; Dorner, F.
  Mol. Microbiol. 18, 257-269, 1995
A;Title: Evidence for lateral transfer and recombination in OspC variation in Lyme disea
A;Accession: S70281; MUID:96296448
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 19-210 <LIV>
A;Cross-references: EMBL:L42887; NID:g858715; PIDN:AAB36995.1; PID:g1695212
A;Gene: ospC
C;Superfamily: Lyme disease spirochete surface protein C
```

```
Query Match 34.7%;  Score 623.5;  DB 2;  Length 210;
Best Local Similarity 69.5%;  Pred. No. 1.3e-22;

Matches 136;  Conservative 20;  Mismatches 26;  Indels 2;  Gaps 2;

Qy    5 NSGKDGNASANSADESVGPNLTEISKKITESNAVVLAVKEVEVETLLASIDELATKAIGKK 64
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 NSGKGNTSANSADESVGPNLTEISKKITESNAVVLAVKEIETLLASIDELATKAIGKK 60

Qy   65 I-GNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKN-EELKEKIDTAKQCSTEFTNKLK 122
          | |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db   61 IQQNNGLAVEAGHNGTLLAGAYVTISKLITQKLDGLKNSEKLKEKIENAKKCSEDFTKLE 120

Qy  123 SEHAVLGLDNLTDDNAQRAILKKHANKDGAAELEKLFKAVENLSKAAQDTLKNAVKELT 182
          ||:|:|||||||:|:||||| ||||||| |||||||||||||||||:||||| |:||||
Db  121 GEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLAKAAKEMLANSVKELT 180

Qy  183 SPIV 186
          ||||
Db  181 SPIV 184
```

RESULT 7

```
S69917
outer surface protein C precursor - Lyme disease spirochete (strain B. pacificus)
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
A;Variety: strain B. pacificus
A;Residues: 1-209 <JAU>
A;Cross-references: EMBL:X83555
A;Experimental source: strain B. pacificus
R;Roessler, D.
A;Reference number: S72679
  submitted to the EMBL Data Library, January 1995
A;Accession: S72679
A;Molecule type: DNA
A;Residues: 1-50,'E',52-61,'I',63-117,'D',119-122,'V',124-173,'D',175-176,'D',178-209
R;Jauris-Helpke, S.; Liegl, G.; Preac-Mursic, V.; Roessler, D.; Schwab, E.; Soutschek
  J. Clin. Microbiol. 33, 1860-1866, 1995
A;Title: Molecular analysis of genes encoding outer surface protein C (OspC) of Borre
A;Reference number: I40047; MUID:95395018
A;Accession: S69917
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-209 <JAU>
A;Cross-references: EMBL:X83555; NID:g872019; PIDN:CAA58545.1; PID:g872020
A;Superfamily: Lyme disease spirochete surface protein C
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-209/Product: outer surface protein C #status predicted <MAT>
```

```
Query Match 34.6%;  Score 622;  DB 2;  Length 209;
Best Local Similarity 69.9%;  Pred. No. 1.6e-22;

Matches 130;  Conservative 24;  Mismatches 32;  Indels 1;  Gaps 1;

Qy    1 MACNNSGKDGNASANSADESVGPNLTEISKKITESNAVVLAVKEVEVETLLASIDELATKA 60
          :||||||||| |||||||||||||||||||||||||||||||||||||:||:|||| |
Db   17 ISCNNSGKDGNTSANSADESVGPNLTEISKKITDSNAVVLAVKEIETLLASIDELAAKA 76

Qy   61 IGKKI-GNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTN 119
          |:|| | |||||||||||||||||||||||||||||||||| |||||||||||||||||
Db   77 IGKKIHQNNGLDTENNHINGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSEFFTN 136

Qy  120 KLKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVK 179
          |||:||:|||||||:|:||||||||||:|| |||||||||||||||||||||:||| |
Db  137 KLKEKHFDLGKEGVTDADAKEAILKTNGTKTKGAAELGKLFESVEVLSKAAKEMLANSVK 196

Qy  180 ELTSPIV 186
          ||||:|:|
Db  197 ELTSPVV 203
```

RESULT 8

```
S69927
outer surface protein C precursor - Lyme disease spirochete (strain PKa)
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
```

```
Query Match 34.6%;  Score 622;  DB 2;  Length 209;
Best Local Similarity 69.9%;  Pred. No. 1.6e-22;

Matches 130;  Conservative 24;  Mismatches 32;  Indels 0;  Gaps 0;

Qy    1 MACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVEVETLLASIDELATKA 60
          :||||||||| |||||||| ||||||||||||||||||||||||||||:||:|||||||
Db   17 ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVVLAVKEIETLLASIDELATKA 76

Qy   61 IGKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNK 120
          ||||||||||||| |||||||||||||||:| |:|||||:||:||||||:||| |||||
Db   77 IGQKIDANGLGVQANQNGSLLAGAYAISTLITQKLSALNSEELKEKIAKVKKCSEDFTNK 136

Qy  121 LKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKE 180
          |:||:|:|||||:|:||||||| |||||: ||||||||:|:|:::|:|:::::|:|||:
Db  137 LKNGNAQLGLAAATDDNAKAAILKTNGTNDKGAKELKELSESVESLVKAAQVMLTNSVKE 196

Qy  181 LTSPIV 186
          |||:|:|
Db  197 LTSPVV 202
```

A;Variety: strain Pka
C;Date: 15-Feb-1997 #sequence_revision 27-Feb-1997 #text_change 26-May-2000
C;Accession: S69927; S72669
R;Jauris-Helpke, S.; Liegl, G.; Preac-Mursic, V.; Roessler, D.; Schwab, E.; Soutschek,
J. Clin. Microbiol. 33, 1860-1866, 1995
A;Title: Molecular analysis of genes encoding outer surface protein C (OspC) of Borrelia
A;Reference number: I40047; MUID:95395018
A;Accession: S69927
A;Molecule type: DNA
A;Status: nucleic acid sequence not shown
A;Residues: 1-210 <JAU>
A;Cross-references: EMBL:X69589
A;Experimental source: strain Pka
R;Jauris, S.
A;Reference number: S72669
submitted to the EMBL Data Library, February 1994
A;Accession: S72669
lecule type: DNA
sidues: 1-124,'D',126-139,'E',141-210 <JAU>
Matches 129; Conservative 25; Mismatches 32; Indels 1; Gaps 1;
Best Local Similarity 69.0%; Pred. No. 2.1e-22;
Query Match 34.5%; Score 619.5; DB 2; Length 210;
C;Superfamily: Lyme disease spirochete surface protein c
C;Genetics:
A;Gene: ospc

Qy   1 MACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKA 60
        ::|||||||| ||||||||||||||||||||||||||||||||:|||||||||||||||||
Db   17 ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVVLAVKEVEVALLSSIDEIAAKA 76

Qy   61 IGKKI-GNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTN 119
        ||||| |||||||:::|||||||||||||:|| :|:| ||||||||||:|||||| |||
Db   77 IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIEAAKKCSEIFTN 136

Qy 120 KLIKSEHAVILGLDNLTTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVK 179
        ||| || |:|| :|| ||||||||||||||||||||||||||||:||||||||||||:|||
Db 137 KLKDKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 196

Qy 180 ELTSPIV 186
        |||||:|
Db 197 ELTSPVV 203

LT   9
86
outer surface protein C - Lyme disease spirochete
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 26-May-2000
C;Accession: S70286
R;Livey, I.; Gibbs, C.P.; Schuster, R.; Dorner, F.
Mol. Microbiol. 18, 257-269, 1995
A;Title: Evidence for lateral transfer and recombination in OspC variation in Lyme disea
A;Reference number: S70255; MUID:96296448
A;Accession: S70286
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-193 <LIV>
A;Experimental source: strain 26815
C;Experimental source: strain 26815
C;Cross-references: EMBL:L42897; NID:g858728; PIDN:AAB37006.1; PID:g1695222
C;Genetics:
C;Gene: ospc
C;Superfamily: Lyme disease spirochete surface protein c

Query Match 34.4%; Score 619; DB 2; Length 193;
Best Local Similarity 71.1%; Pred. No. 2e-22;
Matches 133; Conservative 22; Mismatches 28; Indels 4; Gaps 3;

Qy   3 CNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKAIG 62
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

A:Experimental source: strain PLJ7
R:Roessler, D.
submitted to the EMBL Data Library, September 1994
A:Reference number: S72675
A:Accession: S72675
A:Molecule type: DNA
A:Residues: 1-125,'V',127-154,'E',156-212 <ROE>
A:Cross-references: EMBL:X81523; NID:g804956; PIDN:CAA57243.1; PID:g804957
C:Genetics:
A:Gene: ospC
C:Superfamily: Lyme disease spirochete surface protein C

```
Query Match     34.1%;  Score 612.5;  DB 2;  Length 212;
Best Local Similarity    69.3%;  Pred. No. 4.4e-22;
Matches 131;  Conservative  27;  Mismatches  28;  Indels  3;  Gaps  3;

Qy    1 MACNNSGKDG-NASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATK 59
         ::||||||-|-:||-||||||||||||||||||||||||||:||||:::||||:|||||:
Db   17 ISCNNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETLVSSIDELAAQ 76

Qy   60 AIGKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVL-KNE-ELKEKIDTAKQCSTEF 117
         |||||:||||||||||||||||||||||||||||| :|| ||| ::||||||||||:||:
Db   77 AIGKKIQNNGLTAEQNQNGSLLAGAYAISALITKKLDELTKNSGELKGEIEKAKKCSEEF 136

Qy  118 TNKLKSEHAVLGLDNLITDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNA 177
         |||| |||||| |||||||||||||||||||||:|||||||||||||||||||||::|| :
Db  137 TNKLKGGHAELGLAAATDNAKKAILKTNGTKDKGAEELEKLFKSVESLAKAAKESLTNS 196

Qy  178 VKELTSPIV 186
         |||:|:|
Db  197 VKELTNPVV 205
```

RESULT 12
S70268
outer surface protein C - Lyme disease spirochete
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C:Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 26-May-2000
C:Accession: S70268
R:Livey, I.; Gibbs, C.P.; Schuster, R.; Dorner, F.
Mol. Microbiol. 18, 257-269, 1995
A:Title: Evidence for lateral transfer and recombination in OspC variation in Lyme disea
A:Reference number: S70255; MUID:96296448
A:Accession: S70268
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-194 <LIV>
A:Cross-references: EMBL:L42888; NID:g858716; PIDN:AAB36996.1; PID:g1695213
...netics:
A:Gene: ospC
...xperimental source: strain H9
C:Superfamily: Lyme disease spirochete surface protein C

```
Query Match     33.7%;  Score 605.5;  DB 2;  Length 194;
Best Local Similarity    69.0%;  Pred. No. 8.3e-22;
Matches 129;  Conservative  27;  Mismatches  28;  Indels  3;  Gaps  3;

Qy    3 CNNSGKDG-NASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKAI 61
         ||||||:|-||||||||||:|||||||||||||:::||:|||||||||||||:|||||| ||
Db    1 CNNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETLVSSIDELAAQAI 60

Qy   62 GKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVL-KNE-ELKEKIDTAKQCSTEFTN 119
         |||:|||||||||||||||||||||||||||| :|| ||| ::||||||||||:||:
Db   61 GKKIQNNGLTAEQNQNGSLLAGAYAISALITKKLDELTKNSGELKGEIEKAKKCSEEFTN 120

Qy  120 KLKSEHAVLGLDNLITDDNAQRAILKKHANKDKGAAELEKLFRAVENLSKAAQDTLKNAVK 179
         |||||||| |||||||||||||||||||||:|||||||||||:||||||||||::|| 
Db  121 KLKGGHAELGLAAATDNAKKAILKTNGTKDKGAEELEKLFKSVESLAKAAKESLTNSVK 180

Qy  180 ELTSPIV 186
         |||:|:|
Db  181 ELTNPVV 187
```

Db 181 ELTNPVV 187

RESULT 13
I40104
outer surface protein C - Lyme disease spirochete (strain 272) (fragment)
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C:Status: preliminary; translated from GB/EMBL/DDBJ
C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 26-May-2000
C:Accession: I40104; S54184
C:Genetics:
A:Gene: ospC
R:Theisen, M.; Borre, M.; Mathiesen, M.J.; Mikkelsen, B.; Lebech, A.M.; Hansen, K.
J. Bacteriol. 177, 3036-3044, 1995
A:Title: Evolution of the Borrelia burgdorferi outer surface protein OspC.
A:Reference number: I40104; MUID:95286481
A:Accession: I40104
A:Molecule type: DNA
A:Residues: 1-178 <RES>
A:Cross-references: EMBL:X84785; NID:g793787; PIDN:CAA59256.1; PID:g793788
C:Superfamily: Lyme disease spirochete surface protein C

```
Query Match     33.7%;  Score 605;  DB 2;  Length 178;
Best Local Similarity    73.0%;  Pred. No. 7.9e-22;
Matches 130;  Conservative  19;  Mismatches  27;  Indels  2;  Gaps  2;

Qy   11 NASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKAIGKKI-GNNG 69
         ||||||||||:|||||||||||||:::||:||||||||||||:|||||:|||||:-|||
Db    1 NTSANSADESVKGPNLTEISKKITDSNAFVLAVKEVETLVSSIDELATKAIGKKIQQNGG 60

Qy   70 LEANQSKNTSLLSGAYAISDLIAEKLNVLKN-EELKEKIDTAKQCSTEFTNKLKSEHAVL 128
         |  :|| ||||||||||||||||| |:|| ||| ::||||||||||:||:|||||:||
Db   61 LAVEAGHNGTSLAGAYAISKLITQKLDGLKNSEKLKEKIENAKRCSEDFTKKLSEHAVL 120

Qy  129 GLDNLITDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKELTSPIV 186
         |||||||||||||||||||||:|||||||||||||||||| :|||||||:||||||||
Db  121 GIENVTDENAKKAILTDAARDKGAAELEKLFKAVENLPRAAKEMLANSVKELTSPIV 178
```

RESULT 14
I40125
outer surface protein C - Lyme disease spirochete (strain MUL) (fragment)
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 26-May-2000
C:Accession: I40125; S54195
R:Theisen, M.; Borre, M.; Mathiesen, M.J.; Mikkelsen, B.; Lebech, A.M.; Hansen, K.
J. Bacteriol. 177, 3036-3044, 1995
A:Title: Evolution of the Borrelia burgdorferi outer surface protein OspC.
A:Reference number: I40125; MUID:95286481
A:Accession: I40125
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-178 <RES>
A:Cross-references: EMBL:X84779; NID:g793817; PIDN:CAA59250.1; PID:g793818
C:Superfamily: Lyme disease spirochete surface protein C

```
Query Match     33.7%;  Score 605;  DB 2;  Length 178;
Best Local Similarity    73.0%;  Pred. No. 7.9e-22;
Matches 130;  Conservative  19;  Mismatches  27;  Indels  2;  Gaps  2;

Qy   11 NASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKAIGKKI-GNNG 69
         |||||||||:|||||||||||||:::||:||||||||||||:|||||:|||||:-|||
Db    1 NTSANSADESVKGPNLTEISKKITDSNAFVLAVKEVETLVSSIDELATKAIGKKIQQNGG 60

Qy   70 LEANQSKNTSLLSGAYAISDLIAEKLNVLKN-EELKEKIDTAKQCSTEFTNKLKSEHAVL 128
         |  :|| ||||||||||||||||| |:|| ||| ::||||||||||:||:|||||:||
Db   61 LAVEAGHNGTSLAGAYTISKLITQKLDGLKNSEKLKEKIENAKRCSEDFTKKLEGEHAQL 120

Qy  129 GLDNLITDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKELTSPIV 186
         ||||||||||||||||||||||:|||||||||||||||||| :||||||||||||||||
Db  121 GIENVTDENAKKAILTDAARDKGAAELEKLFKAVENLPKRAAKEMLANSVKELTSPIV 178
```

RESULT 15

I40144

outer surface protein C precursor - Lyme disease spirochete
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 21-Jul-2000
C;Accession: I40144; S70282
R;Stevenson, B.; Barthold, S.W.
FEMS Microbiol. Lett. 124, 367-372, 1994
A;Title: Expression and sequence of outer surface protein C among North American isolate
A;Reference number: I40143; MUID:95154673
A;Accession: I40144
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-210 <RES>
A;Cross-references: EMBL:U04281; NID:g434663; PIDN:AAC43297.1; PID:g434664
R;vey, I.; Gibbs, C.P.; Schuster, R.; Dorner, F.
Microbiol. 18, 257-269, 1995
A;Title: Evidence for lateral transfer and recombination in OspC variation in Lyme disea
A;Reference number: S70255; MUID:96296448
A;Accession: S70282
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 19-210 <LIV>
A;Cross-references: EMBL:L42893; NID:g858721; PIDN:AAB37001.1; PID:g1695218
A;Experimental source: strain 297
C;Genetics:
A;Gene: OspC
C;Superfamily: Lyme disease spirochete surface protein C

Query Match   33.6%;   Score 603.5; DB 2;   Length 210;
Best Local Similarity   68.6%;   Pred. No. 1.1e-21;
Matches 129; Conservative 26; Mismatches 30; Indels 3; Gaps 3;

Qy   1 MACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKA 60
         ::||||||||| |||||||||||||||||||||||||||||||||||||||||| ||
Db   17 ISCNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLTSIDELA-KA 75

Qy   61 IGKKIGNN-GLEANQSKNTSLLSGAYAISDLIAEKLNVLKNE-ELKEKIDTAKQCSTEFT 118
         ||||| |: |||||||: ||:||||||||||||: ||:|| |||
Db   76 IGKKIKNDVSLDNEADHNGSLISGAYLISTLITKKISAIKDSGELKAEIEKAKKCSEEFT 135

Qy   119 NKLKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAV 178
         ||| || | | | |:||| | | ||| ::||||||||::|||||||||::::|:|
Db   136 AKLKGEHTDLGKEGYVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLSKAAKEMLTNSV 195

Qy   179 KELTSPIV 186
         ||||||:|
Db   196 KELTSPVV 203

Search completed: March 18, 2002, 09:56:48
Job time: 371 sec

OM protein - protein search, using sw model

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

Run on: March 18, 2002, 10:11:52 ; Search time 39.62 seconds
                 (without alignments)
                 340.552 Million cell updates/sec

Title: US-09-596-746A-38
Perfect score: 1797
Sequence: 1 MACNNSGKDGNASANSADES.........ENLSKAAQDTLKNAVKELTS 368

Scoring table: BLOSUM62
               Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059
Maximum DB seq length: 2000000000
Minimum DB seq length: 0

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID		Description
1	623.5	34.7	210	1	OSC1_BORBU	Q07337	borrelia bu
2	588.5	32.7	212	1	OSC2_BORBU	Q08337	borrelia bu
3	346	19.3	214	1	VM24_BORHE	P3278	borrelia he
4	311.5	17.3	215	1	VM03_BORHE	P3278	borrelia he
5	172.5	9.6	1790	1	USO1_YEAST	P25386	saccharomyc
6	166.5	9.3	944	1	NUF1_YEAST	P32380	saccharomyc
7	163.5	9.1	1167	1	CAGA_HELPJ	Q9it1	helicobacte
8	162.5	9.0	2017	1	MYSN_DROME	Q99923	drosophila
9	159.5	8.9	1875	1	MLP1_YEAST	Q02455	saccharomyc
10	158.5	8.8	705	1	YN9J_CAEEL	P34562	caenorhabdi
11	157.5	8.8	1251	1	RBP2_PLAVB	Q00799	plasmodium
12	157	8.7	1639	1	LMG1_DROME	P15215	drosophila
13	156.5	8.7	1013	1	A60D_DROME	P91927	drosophila
14	154	8.6	1526	1	MYS2_SCHPO	Q9us16	schizosacch
15	153	8.5	548	1	CEAK_ECOLI	Q47502	escherichia
16	152.5	8.5	1191	1	MST2_DROHY	Q08696	drosophila
17	151.5	8.4	1966	1	MYSB_CAEEL	P02566	caenorhabdi
18	150	8.3	1147	1	CGA1_HELPY	P80200	helicobacte
19	149.5	8.3	1786	1	LMB1_HUMAN	P07942	homo sapien
20	149.5	8.3	2116	1	MYS2_DICDI	P08799	dictyosteli
21	148.5	8.3	1957	1	YD86_SCHPO	Q10411	schizosacch
22	148.5	8.3	2249	1	TPR_HUMAN	P12270	homo sapien
23	148	8.2	1433	1	REST_CHICK	O42184	gallus gall
24	148	8.2	1701	1	MSP1_PLAFF	P13819	plasmodium
25	147.5	8.2	1102	1	MYSC_CHICK	P29616	gallus gall
26	147.5	8.2	1182	1	CGA2_HELPY	P55746	helicobacte
27	147	8.2	845	1	SCP1_MESAU	Q60563	mesocricetu
28	147	8.2	918	1	YMJB_CAEEL	P34487	caenorhabdi
29	147	8.2	1427	1	REST_HUMAN	Q30622	homo sapien
30	146.5	8.2	1637	1	MRSP_STAAU	P80544	staphylococ
31	146.5	8.2	1786	1	LMB1_MOUSE	P02469	mus musculu
32	145.5	8.1	1938	1	MYH4_RABIT	Q28641	oryctolagus
33	145	8.1	539	1	MYS3_HYDAT	P39922	hydra atten
34	144.5	8.0	1044	1	YAF3_SCHPO	Q09857	schizosacch
35	144	8.0	483	1	M6_STRPY	P08089	streptococc
36	143.5	8.0	1509	1	MYSN_ACACA	P05659	acanthamoeb
37	143	8.0	364	1	VM21_BORHE	P21875	borrelia he
38	143	8.0	1186	1	CAGA_HELPY	P55980	helicobacte
39	143	8.0	1727	1	ALM1_SCHPO	Q9utk5	schizosacch
40	142.5	7.9	1727	1	GCG4_HUMAN	Q13439	homo sapien
41	142	7.9	522	1	CEA1_ECOLI	P02978	escherichia
42	142	7.9	1186	1	SMC_BACSU	P51834	bacillus su
43	141.5	7.9	564	1	M12_STRPY	P19401	streptococc
44	141	7.8	276	1	SFAS_CHLRE	Q39618	chlamydomon
45	141	7.8	1701	1	MSP1_PLAFM	P08569	plasmodium

ALIGNMENTS

RESULT 1
OSC1_BORBU

ID   OSC1_BORBU   STANDARD;   PRT;   210 AA.
AC   Q07337;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   OUTER SURFACE PROTEIN C PRECURSOR (PC).
DE   OSPC OR BBB19.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OG   Plasmid lp54.
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 35210 / B31;
RX   MEDLINE=93268136; PubMed=8098841;
RA   Jauris-Heipke S., Fuchs R., Motz M., Preac-Mursic V., Schwab E.,
RA   Will G., Wilske B.;
RT   "Genetic heterogeneity of the genes coding for the outer surface
RT   protein C (OspC) and the flagellin of Borrelia burgdorferi.";
RL   Med. Microbiol. Immunol. 182:37-50(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 35210 / B31;
RX   MEDLINE=93293332; PubMed=8478108;
RA   Wilske B., Preace-Mursic V., Jauris S., Pradel I., Soutschek E.,
RA   Schwab E., Wanner G.;
RT   "Immunological and molecular polymorphisms of OspC, an immunodominant
RT   major outer surface protein of Borrelia burgdorferi.";
RL   Infect. Immun. 61:2182-2191(1993).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 35210 / B31;
RX   MEDLINE=94041630; PubMed=8225587;
RA   Padula S.J., Sampieri A., Dias F., Szczepanski A., Ryan R.W.;
RT   "Molecular characterization and expression of p23 (OspC) from a North
RT   American strain of Borrelia burgdorferi.";
RL   Infect. Immun. 61:5097-5105(1993).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 35210 / B31;
RX   MEDLINE=96025162; PubMed=7494039;
RA   Fukunaga M., Hamase A.;
RT   "Outer surface protein C gene sequence analysis of Borrelia
RT   burgdorferi sensu lato isolates from Japan.";
RL   J. Clin. Microbiol. 33:2415-2420(1995).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 35210 / B31;
RX   MEDLINE=98065943; PubMed=9403685;
RA   Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA   Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
RA   Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
RA   Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,
RA   van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,

RA Utterback T., Watthey L., McDonald L., Artiach P., Bowman C.,
RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
RA Smith H.O., Venter J.C.;
RT "Genomic sequence of a Lyme disease spirochaete, Borrelia
RT burgdorferi.";
RL Nature 390:580-586(1997).
CC   [2]
RN   FUNCTION: NOT KNOWN; MAJOR IMMUNODOMINANT PROTEIN.
CC -!- SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID
CC   ANCHOR.
CC ---
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC ---
DR   EMBL; X69596; CAA49306.1; -.
DR   EMBL; U01894; AAA16058.1; -.
DR   EMBL; D49497; BAA08457.1; -.
DR   EMBL; AE000792; AAC66329.1; -.
DR   TIGR; BB0119; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
DR   PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KW   Outer membrane; Lipoprotein; Signal; Plasmid; Antigen;
KW   Complete proteome.
FT   SIGNAL   1   18   BY SIMILARITY.
FT   CHAIN   19   210   OUTER SURFACE PROTEIN C.
FT   LIPID   19   19   N-ACYL DIGLYCERIDE (BY SIMILARITY).
SQ   SEQUENCE   210 AA; 22340 MW; 7A4FC978F91777BF CRC64;

Query Match   34.7%; Score 623.5; DB 1; Length 210;
Best Local Similarity   69.5%; Conservative   24; Mismatches   32; Indels   1; Gaps   1;
Matches   130;   Pred. No. 5.3e-22;

Qy   1 MACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKA 60
          :::||||||| |||||||||:|:||||||||||||:|||||||| ||:|||||::|:||
Db   17 ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA 76

Qy   61 IGKKI-GNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTN 119
          ||||| ||||||||||||| |||||||||| || | | | | | |||||||||| ||
Db   77 IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN 136

Qy   120 KLKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVK 179
          |||:|| |:|| ||| || :||||||| |:|| ||:|||||||||||||::|||||||:|
Db   137 KLKEKHTDLGKEGVTDADAKEAILKTNGIKTKGAEELGKLFESVEVLSKAAKEMLANSVK 196

Qy   180 ELISPIV 186
          |||||:|
Db   197 ELISPVV 203

RA Soutschek E.;
RT "Molecular analysis and expression of a Borrelia burgdorferi gene
RT encoding a 22 kDa protein (pC) in Escherichia coli.";
RL Mol. Microbiol. 6:503-509(1992).
RN   [2]
RP   SEQUENCE OF 1-205 FROM N.A.
RC   STRAIN=DK26;
RX   MEDLINE=94075528; PubMed=8253951;
RA   Theisen M., Frederiksen B., Lebech A.M., Vuust J., Hansen K.;
RT   "Polymorphism in ospC gene of Borrelia burgdorferi and
RT   immunoreactivity of OspC protein: implications for taxonomy and for
RT   use of OspC protein as a diagnostic antigen.";
RL   J. Clin. Microbiol. 31:2570-2576(1993).
CC -!- FUNCTION: NOT KNOWN; MAJOR IMMUNODOMINANT PROTEIN.
CC -!- SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID
CC   ANCHOR.
CC ---
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC ---
DR   EMBL; X62162; CAA44093.1; -.
DR   EMBL; X73624; CAA52003.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
KW   Outer membrane; Lipoprotein; Signal; Plasmid; Antigen.
FT   SIGNAL   1   18   BY SIMILARITY.
FT   CHAIN   19   212   OUTER SURFACE PROTEIN C.
FT   LIPID   19   19   N-ACYL DIGLYCERIDE (BY SIMILARITY).
SQ   SEQUENCE   212 AA; 22499 MW; C206C231F8F2E7D4 CRC64;

Query Match   32.7%; Score 588.5; DB 1; Length 212;
Best Local Similarity   68.8%; Conservative   20; Mismatches   36; Indels   3; Gaps   3;
Matches   130;   Pred. No. 1.9e-20;

Qy   1 MACNNSGKDG-NASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATK 59
          :::||||||| ||||||||||||||||||||||:||||| ||||||||:|||||||||
Db   17 ISCNNSGKGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETLVLSIDELAKK 76

Qy   60 AIGKKI-GNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKN-EELKEKIDTAKQCSTEF 117
          |||||| || ||||||||||| ||||||||||||| ||||| ||||||||| ||||||
Db   77 AIGQKIDNNNGLAALNNQNGSLLAGAYAISTLITEKLSKLRNLEELKTEIAKAAKKCSEEF 136

Qy   118 TNKLKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNA 177
          |||||:|||| ||| |||| :||||||||| ||:|| ||:|||||||||||||::|| |
Db   137 TNKLKSGHADLGKQDATDDHAKAAILKTHATTDKGAKEEKDLFESVEGLLKAAQVALTNS 196

Qy   178 VKELISPIV 186
          |||||||:|
Db   197 VKELISPVV 205

ID   OSC2_BORBU   STANDARD;   PRT;   212 AA.
AC   Q08137;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   OUTER SURFACE PROTEIN C PRECURSOR (PC).
GN   OSPC.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
RC   STRAIN=PKO;
RX   MEDLINE=92219995; PubMed=1560779;
RA   Fuchs R., Jauris S., Lottspeich F., Preac-Mursic V., Wilske B.,

RESULT 2
OSC2_BORBU

ID   VM24_BORHE   STANDARD;   PRT;   214 AA.
AC   P32778;
DT   01-OCT-1993 (Rel. 27, Created)
DT   01-OCT-1993 (Rel. 27, Last sequence update)
DT   01-FEB-1994 (Rel. 28, Last annotation update)
DE   VARIABLE MAJOR OUTER MEMBRANE LIPOPROTEIN 24 PRECURSOR.
GN   VMP24.
OS   Borrelia hermsii.
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=140;
RP   [1]
RP   SEQUENCE FROM N.A.

RESULT 3
VM24_BORHE

RC   STRAIN=SSP, HS1 SEROTYPE 24;
RX   MEDLINE=93133110; PubMed=1484486;
RA   Restrepo B.I., Kitten T., Carter C.J., Infante D., Barbour A.G.;
RT   "Subtelomeric expression regions of Borrelia hermsii linear plasmids
RT   are highly polymorphic.";
RL   Mol. Microbiol. 6:3299-3311(1992).
CC   -!- FUNCTION: SERVES TO AVOID THE HOST IMMUNE RESPONSE BY CHANGING
CC   FROM ONE SURFACE EXPOSED VMP TO ANOTHER.
CC   -!- SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID
CC   ANCHOR.
CC   -!- SIMILARITY: STRONG, TO VMP3.
CC   ---
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---
DR   EMBL; L04786; AAA22964.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
DR   PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KW   Outer membrane; Lipoprotein; Signal; Plasmid.
FT   SIGNAL   1   18   PROBABLE.
FT   CHAIN   19   214   VARIABLE MAJOR OUTER MEMBRANE
FT   LIPOPROTEIN 24.
FT   LIPID   19   19   N-ACYL DIGLYCERIDE (PROBABLE).
SQ   SEQUENCE   214 AA; 22541 MW; F1583F510246F7C7 CRC64;

Query Match   19.3%; Score 346; DB 1; Length 214;
Best Local Similarity   42.1%;
Matches   82; Conservative 40; Mismatches 57; Indels 16; Gaps   5;

Qy   1 MACNNSGKDGNASANSADESVKGP----NLTEISKKITESNAVVLAVKEVETLLASIDEL 56
             :||:|.| | ||::||||||| |:||||||||:||||||||||||||:|||||
Db   17 MSCNNGGPE----LKSDEVAKSDGTVLDLAKVSKKIKEASAFAASVKEVETLVKSVDEL 71

Qy   57 ATKAIGKKIGNN-GLEANQSKNTSLLSGAYAISDLIAEKLNVLK-----NEELKEKIDTA 110
             ||||.|.| |: |:|||:|| |:||| |||||||||||||||| ||||||||:|
Db   72 A-KAIGKKIKNDGGLDTEAGQNGSLIAGVHSVVSAVKIKVGALETTSGISNELKTKITEV 130

Qy   111 KQCSTEFTNKLKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAA 170
             | .| :|:||||:||||||||:|:|| |||||||:|||||||:|| |||||:|||
Db   131 KSKAEAFLNKLRDGHTELGKKDASDDDFKKAIKKDNSDKTKGASELEALNTAVDALLKAA 190

Qy   171 QDTLKNAVKELTSPI 185
             .||||||:|||:|
Db   191 EGEVEAAIKELTAPV 205

RT   are highly polymorphic.";
RL   Mol. Microbiol. 6:3299-3311(1992).
CC   -!- FUNCTION: SERVES TO AVOID THE HOST IMMUNE RESPONSE BY CHANGING
CC   FROM ONE SURFACE EXPOSED VMP TO ANOTHER.
CC   -!- SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID
CC   ANCHOR.
CC   -!- SIMILARITY: STRONG, TO VMP24.
CC   ---
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---
DR   EMBL; L04789; AAA22967.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
DR   PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KW   Outer membrane; Lipoprotein; Signal; Plasmid.
FT   SIGNAL   1   18   PROBABLE.
FT   CHAIN   19   215   VARIABLE MAJOR OUTER MEMBRANE
FT   LIPOPROTEIN 3.
FT   LIPID   19   19   N-ACYL DIGLYCERIDE (PROBABLE).
SQ   SEQUENCE   215 AA; 23139 MW; 684C74D35F87C771 CRC64;

Query Match   17.3%; Score 311.5; DB 1; Length 215;
Best Local Similarity   39.8%;
Matches   78; Conservative 36; Mismatches 65; Indels 17; Gaps   5;

Qy   1 MACNNSGKDGNASANSADESVKGP----NLTEISKKITESNAVVLAVKEVETLLASIDEL 56
             :||:|.| | ||::||||||| |:||||||||:||||||||||||||:|||||
Db   17 MSCNNGGPE----LKSDEVAKSDGTVLDLAKRISKKIKDASDFAASVKEVHTLVKSIDEL 71

Qy   57 ATKAIGKKI-GNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLK-----NEELKEKIDT 109
             ||||.|.| |:|||:||| |:||| |||||||||||||||| ||||||||:|
Db   72 A-KAIGKKIKNDNSNFEDENDHNGSLIAGVFQVILTVKAKLTSLEQIIGISDELKTEVGM 130

Qy   110 AKQCSTEFTNKLKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKA 169
             | .| :|:||||:||||||||:|:|| |||||||:|||||||:|| |||||:||
Db   131 VKKESEAFVTQVKSKHTDLAREGVTDAHAKSAILVTDGTKDKGAAELIKLNTAIDELLKA 190

Qy   170 AQDTLKNAVKELTSPI 185
             .||||||:|||:|
Db   191 ANDAVETVIKELTASV 206

RESULT 5
USO1_YEAST
ID   USO1_YEAST   STANDARD;   PRT;   1790 AA.
AC   P25386;
DT   01-MAY-1992 (Rel. 22, Created)
DT   01-MAY-1992 (Rel. 22, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   INTRACELLULAR PROTEIN TRANSPORT PROTEIN USO1.
GN   USO1 OR INT1 OR YDL058W.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=X2180-1A;
RX   MEDLINE=91185402; PubMed=2010462;
RA   Nakajima H., Hirata A., Ogawa Y., Yonehara T., Yoda K.,
RA   Yamasaki M.;
RT   "A cytoskeleton-related gene, uso1, is required for intracellular
RT   protein transport in Saccharomyces cerevisiae.";
RL   J. Cell Biol. 113:245-260(1991).
RN   [2]

RESULT 4
VM03_BORHE
ID   VM03_BORHE   STANDARD;   PRT;   215 AA.
AC   O02448;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   01-FEB-1994 (Rel. 28, Last annotation update)
DE   VARIABLE MAJOR OUTER MEMBRANE LIPOPROTEIN 3 PRECURSOR.
GN   VMP3.
OS   Borrelia hermsii.
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=140;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SSP, HS1 SEROTYPE 3;
RX   MEDLINE=93133110; PubMed=1484486;
RA   Restrepo B.I., Kitten T., Carter C.J., Infante D., Barbour A.G.;
RT   "Subtelomeric expression regions of Borrelia hermsii linear plasmids

```
RP   SEQUENCE OF 782-1790 FROM N.A.
RA   Hostetter M.K., Herman D.J., Bendel C.M., McClellan M., Tao N.,
RA   Kendrick K.E.;
RL   Submitted (FEB-1993) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 1-8 FROM N.A.
RA   Bai Y., Symington L.S.;
RL   Submitted (MAY-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: REQUIRED FOR PROTEIN TRANSPORT FROM THE ER TO THE GOLGI
CC       COMPLEX.
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC. ASSOCIATED WITH INTRACELLULAR
CC       MEMBRANES. PROBABLY PRESENT ON VESICLES OPERATIONAL BETWEEN THE
CC       ER AND THE GOLGI COMPLEX.
CC   -!- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, COMPOSED
CC       OF AN HEPTAPEPTIDE REPEAT PATTERN CHARACTERISTIC OF ALPHA-HELICAL
CC       COILED COILS. MAY FORM FILAMENTOUS STRUCTURES IN THE CELL.
CC   -!- SIMILARITY: BELONGS TO THE VDP/USO1/YBL047C FAMILY.
CC   -!-
```

```
DR   EMBL; X54378; CAA38253.1; -.
DR   EMBL; L03188; AAB00143.1; -.
DR   EMBL; U53668; AAB66659.1; -.
DR   PIR; A38455; A38455.
DR   HSSP; P80220; 1DIP.
DR   SGD; S0002216; USO1.
DR   InterPro; IPR002017; Spectrin.
KW   Transport; Protein transport; Golgi stack; Cytoskeleton; Coiled coil.
FT   DOMAIN       1    724    GLOBULAR HEAD.
FT   DOMAIN     725   1790    COILED COIL (POTENTIAL).
FT   DOMAIN     465    487    CHARGED (HYPER-HYDROPHILIC).
FT   DOMAIN     991   1790    DISPENSABLE FOR THE PROTEIN FUNCTION.
FT   DOMAIN    1172   1786    ASP/GLU-RICH (ACIDIC).
FT   CONFLICT   847    847    G -> E (IN REF. 2).
FT   CONFLICT   924    924    E -> K (IN REF. 2).
FT   CONFLICT  1253   1253    V -> I (IN REF. 2).
FT   CONFLICT  1319   1319    V -> V (IN REF. 2).
FT   CONFLICT  1461   1461    N -> S (IN REF. 2).
FT   CONFLICT  1581   1581    G -> S (IN REF. 2).
FT   CONFLICT  1600   1600    I -> V (IN REF. 2).
FT   CONFLICT  1661   1661    R -> S (IN REF. 2).
FT   CONFLICT  1772   1772    D -> DEEDDEE (IN REF. 2).
SQ   SEQUENCE   1790 AA;  206424 MW;  6CE2B216E9FD4818 CRC64;

Query Match  9.68;  Score 172.5;  DB 1;  Length 1790;
Best Local Similarity 22.1%;  Pred. No. 0.6;  Matches 103;  Conservative 66;  Mismatches 152;  Indels 145;  Gaps 20;

Qy     3 CNNSGKDG----------NASANSADESVKGPNLTEISKKI------------TESNAVVLAVK 44
           |||  |   : |   ||| |        | ||  :::  : :  :: |||
Db   929 CNNLSKEKEHISKELEVEKSFQSHDNLV--AKLIEKLKSLANNYKDMQAENESLIKAVE 986

Qy    45 E------VETLLASIDELATKAIGKKIGNNGLEAN--QSKNTSLLSGAYAISDLIAEK 94
Db   987 ESKNESSIQLSNLQNKIDSMSQEKENFQIERGSIEKNIEQLKKT------ISDLEQTK 1038

Qy    95 LNVLKNEE-----------LKEKIDTAKQCSTEFTNKLKS-------EHAVLGLDNL- 133
Db  1039 EEIISKSDSSKDEYESQISLLKEKLETATANDENVNKISELTKTREELEAELAAYKNLK 1098

Qy   134 ---TDDNAQRAI----------------LKKHANKDKG------AAELEKLFKAVENLS 167
Db  1099 NELETKLETSEKALKEVEKENEEHLKEEKIQLEKEATETKQQLNSLRANLESLEKEHEDLA 1158

Qy   168 --KAAQDTLKNAVKELTSPIVHGNNSRKDGNASTNSADESVKGPNLTEI------SKKIT 219
```

```
Db  1159 AQLKKYEEQIANKERQYNEEISQLN-----DEITSTQQENESIKKKN-DELEGEVKAMKST 1213
Qy   220 ESNAVVLAVKEVETLLASIDELATKAIGKKIGNNGLEANQSKNTSLLSGAYAISDLIAEK 279
Db  1214 SEEQSNLKKSEIDALNLQIKELKKK----------------NETNEASLLESIKSVES---- 1255
Qy   280 LNVLKNEELKEKIDTAKQCSTEFTNKLKSEHAVLGLDNLTDDNAQRAILKKHANK------- 334
Db  1256 -ETYKIKELQDECNFKEKEVSELEDKLKASE--------------DKNSKYLELQKESEKIKEEL 1305
Qy   335 DKGAAELEKLFKAVENLSKA------------------AQDTLKNAVKEL 366
Db  1306 DAKTTELKIQLEKITNLSKAKEKSESELSRLKKTSSEERKNAEEQL 1351

RESULT 6
NUF1_YEAST
ID   NUF1_YEAST    STANDARD;    PRT;   944 AA.
AC   P32380;
DT   01-OCT-1993 (Rel. 27, Created)
DT   01-OCT-1993 (Rel. 27, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   NUF1 PROTEIN (SPINDLE POLY BODY SPACER PROTEIN SPC110).
GN   NUF1 OR SPC110 OR YDR356W OR D9476.3.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288C;
RX   MEDLINE=92176232; PubMed=1541631;
RA   Mirzayan C., Copeland C.S., Snyder M.;
RT   "The NUF1 gene encodes an essential coiled-coil related protein that
RT   is a potential component of the yeast nucleoskeleton.";
RL   J. Cell Biol. 116:1319-1332(1992).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94064779; PubMed=7503995;
RA   Kilmartin J.V., Dyos S.L., Kershaw D., Finch J.T.;
RT   "A spacer protein in the Saccharomyces cerevisiae spindle poly body
RT   whose transcript is cell cycle-regulated.";
RL   J. Cell Biol. 123:1175-1184(1993).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288C / AB972;
RA   Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z.,
RA   Favello A., Fulton L., Gattung S., Greco T., Kirsten J.,
RA   Kucaba T., Hillsworth K., Hawkins J., Hillier L., Jier M.,
RA   Johnson D., Johnston L., Langston Y., Latreille P., Le T.,
RA   Mardis E., Menezes S., Miller N., Nhan M., Pauley A., Peluso D.,
RA   Rifken L., Riles L., Taich A., Trevaskis E., Vignati D.,
RA   Wilcox L., Wohldman P., Vaudin M., Wilson R., Waterston R.;
RL   Submitted (JUN-1995) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: ESSENTIAL COMPONENT OF THE NUCLEOSKELETON. POTENTIAL
CC       ROLE IN CROSSLINKING FILAMENTS OR ANCHORING OTHER MOLECULES. IT
CC       IS ESSENTIAL FOR GROWTH.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR; TIGHTLY ASSOCIATED WITH THE
CC       NUCLEOLUS. IT IS PRESENT IN A GRANULAR PATTERN THAT EXCLUDES THE
CC       NUCLEOLUS.
CC   -!- PTM: MAY BE REGULATED BY PHOSPHORYLATION EVENTS.
CC   -!-
```

```
DR   EMBL; Z11582; CAA77668.1; -.
DR   EMBL; X73297; CAA51733.1; -.
DR   EMBL; U28372; AAB64791.1; -.
```

RESULT 7

```
ID   CAGA_HELPJ     STANDARD;      PRT;   1167 AA.
AC   Q9ZLT1;
DT   20-AUG-2001 (Rel. 40, Created)
DT   20-AUG-2001 (Rel. 40, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   CYTOTOXICITY ASSOCIATED IMMUNODOMINANT ANTIGEN (120 KDA PROTEIN)
DE   (CAG PATHOGENICITY ISLAND PROTEIN 26).
GN   CAGA OR CAI OR CAG26 OR JHP0495.
OS   Helicobacter pylori J99 (Campylobacter pylori J99).
OC   Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC   Helicobacter.
     NCBI_TaxID=85963;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99120557; PubMed=9923682;
RA   Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
RA   Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
RA   Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA   Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
RA   Trust T.J.;
RT   "Genomic sequence comparison of two unrelated isolates of the human
RT   gastric pathogen Helicobacter pylori.";
```

Nature 397:176-180(1999).

```
CC   -!- FUNCTION: MAY BE NECESSARY FOR THE TRANSCRIPTION, FOLDING, EXPORT,
CC       OR FUNCTION OF THE CYTOTOXIN.
CC   -------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -------------------------------------------------------------------------
DR   EMBL; AE001483; AAD06073.1; -.
KW   Antigen; Complete proteome.
FT   DOMAIN      246    249       POLY-THR.
FT   DOMAIN      882    889       POLY-ASN.
SQ   SEQUENCE   1167 AA;  129729 MW;  FD5E86B81CEBD0F2 CRC64;
```

```
Query Match     Best Local Similarity   26.2%;
                Matches  102;  Conservative   52;  Mismatches  127;  Indels  109;  Gaps  20;
                                        Score 163.5;  Pred. No. 0.95;  DB 1;  Length 1167;

Qy   24  PNLTEIS----------------KKITESNAVVLAVKEVETLLASIDELATKAIGKKIGNNGLEA  72
             ::                 ::!:  !  ! :!: !!!!   :  :!: !!::!!  !!!::!!
Db  545  PNLNNLAITSVVRQDLEDKLIAKGLSPQEANKLVKDFLSSNKELVGKAL----NFNKAV      599

Qy   73  NQSKNTSLLSGAY-----AISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNKLKSEHAV     127
                 :!  !! !     :: !  :  !!  !!  !::!!::  !! !  !::::
Db  600  AEAKNT----GNYDEVKQAOKDL--EK--SLKKRERLEK-DVAKNLESKSGNKNKMEAK-     649

Qy  128  LGLDNLTDDNAQR-----AILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKELTS     183
              !:!   !  !     ! !:::!:!                   :!  !::!!
Db  650  ------SQANSOKDEIFALINKEANRDARAI------------------AYAQNLKGIKRELSD  689

Qy  184  PIVHGNNSRKDGNASTNSADESVKGPNLTEISKKITESNAVVLAVKE------------VET     233
             !   !:!!  !  :!:  :  :      ::!:         :::!! :        :!
Db  690  KLENINKDLKD--FSKSFDEFKNGKN-KDFSKAEETLKALKGSVKDLGINPEWISKVEN     745

Qy  234  LLASIDELATKAIGKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKID     293
              !!  ! !!  : !:   :!  :: !! !! !                 ::!!
Db  746  LNAALNEF-----------KNGKNKDFSKVTQ------AKSDLENSIKDVIINQKITDKVD     789

Qy  294  TAKQ-------CSTEFTNKLKSEHAVLGLDNLTDDN-AQRAILKKHANKDKGAAELEKLFK     346
              !!        ! :!! !! :!: !  !:!   !  !  ::!::!
Db  790  NLNQAVSVAKATGDFSG---VEQALADLKNFSKEQLAQOAQKNEDENTGKNSA----LYQ     842

Qy  347  AVEN-----------LSKAAQDTLKNAVKEL     366
             !:!            :!:
Db  843  SVKNGVNGTLVGNGLSKAEATTLSKNFSDI     872
```

```
DR   PIR; S26710; S26710.
DR   PIR; S34288; S34288.
     SGD; S002764; NUF1.
KW   Coiled coil; Nuclear protein; Phosphorylation.
FT   DOMAIN      164    791       COILED COIL.
FT   DOMAIN       54     59       NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT   DOMAIN      726    731       NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT   DOMAIN      742    747       NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT   DOMAIN      731    944       NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT   DOMAIN      742    747       ARG/TYR-RICH.
SQ   SEQUENCE   944 AA;  111781 MW;  04FAA074BB8A0BC8 CRC64;
```

```
Query Match     Best Local Similarity    9.3%;
                Matches  100;  Conservative   71;  Mismatches  168;  Indels  89;  Gaps  22;
                                         Score 166.5;  Pred. No. 0.55;  DB 1;  Length 944;

Qy    6  SGKDGNASANSADESVKGPNLTEISKKITESNAV-VLAVKEVETLLASIDE-LATKAIGK      63
                     ::  !:!                :!:  !:!:!       !::
Db   91  SPRSGN----VDKSRKRNLIDDLKKDVPMSQPLKEQEVREHQMKKEREDRALESKLLGK     145

Qy   64  K------IGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKN-EELKEKIDTAKQCS     114
                    :! :  !   :!!  !! :  !!!!!   ::!!
Db  146  RHITYANSDISNKELYINEIKSLK---HEIKELRKEKNDTLNNYDTLEEETDDLKNRL     200

Qy  115  TEFTNKLKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAV-------     167
             !::!!  :    !:!   :  !   :: !::
Db  201  QALEKELDAKNKIVN-SRKVDDHS--GCIEEREQMERKLAELERKLKTVKDQVLELENNS  257

Qy  168  ------KAAQDTLKNAVKELTSPIVHGNNSRKDGNAS------------TNSADE-SVKGPN  210
                  :!  !::!!                               !:!:!
Db  258  DVQSLKRSKEDELKNLMNELNE--LKSNAEEKDTQLEFKNELRKRINELNELKIKSDE     315

Qy  211  L-TEISKKITESNAVVLAVKEVETLL------ASIDELATKAIGKKIGNNGLEAN-QSKN     262
              !::!:!         :::!!      !!  ! !!  : !:   :!  ::
Db  316  MDLQLKQKONESRKLKDELNELETKFSENGSQSSAKENELKMLKNKIAE--LEEEISTKN  373

Qy  263  TSLLSGAYAISDLIAE-----KLNV------LKNEELKEKIDTAKQCSTEFTNKLKSEH     310
              !! !!                !!        !!  !!  !::!!::  !! !  !::
Db  374  SOIIAKEGKLASLMAQLTQLESKLNQRDSQLGSREEELKKTNDKLQK------DIRIAREE  428

Qy  311  AVLGLDNLTD-----------DNAQRAILKKHANK----DKGAAELEKLFKAVENLSKAAQDT  358
             :!  !:!              !  ! !:::!:!                   :!
Db  429  TVSKDERIIDLQKKVKQLENDLFVIKKTHSESKTITDNELESKDKLIKILENDLKVAQEK  488

Qy  359  LKNAVKEL     366
              :!
Db  489  YSKMEKEL     496
```

RESULT 8

```
ID   MYSN_DROME     STANDARD;      PRT;   2017 AA.
AC   Q99323;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   MYOSIN HEAVY CHAIN, NON-MUSCLE (ZIPPER PROTEIN) (MYOSIN II).
GN   ZIP.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
     NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RX   MEDLINE=90349606; PubMed=2117279;
RA   Ketchum A.S., Stewart C.T., Stewart M., Kiehart D.P.;
RT   "Complete sequence of the Drosophila nonmuscle myosin heavy-chain
RT   transcript: conserved sequences in the myosin tail and differential
```

RT   splicing in the 5' untranslated sequence.";
RL   Proc. Natl. Acad. Sci. U.S.A. 87:6316-6320(1990).
CC   -!- FUNCTION: NONMUSCLE MYOSIN APPEARS TO BE RESPONSIBLE FOR
CC   CELLULARIZATION. REQUIRED FOR MORPHOGENESIS AND CYTOKINESIS.
CC   -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
CC   SHORT FORM;
CC   ARE PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 IQ DOMAIN.
CC   ---

DR   EMBL; M35012; AAA28713.1; -.
DR   PIR; A36014; A36014.
DR   HSSP; P08799; 1MND.
DR   FlyBase; FBgn0005634; zip.
DR   InterPro; IPR000048; IQ.
DR   InterPro; IPR002928; Myosin_tail.
DR   InterPro; IPR001609; myosin_head.
DR   Pfam; PF00612; IQ; 1.
DR   Pfam; PF00063; myosin_head; 1.
DR   Pfam; PF01576; Myosin_tail_1; 1.
DR   PRINTS; PR00193; MYOSINHEAVY.
DR   ProDom; PD003355; myosin_head; 1.
DR   SMART; SM00015; IQ; 1.
DR   SMART; SM00242; MYSc; 1.
DR   PROSITE; PS00096; IQ; 1.
KW   Myosin; Alternative splicing; Coiled coil; Actin-binding;
KW   ATP-binding; Calmodulin-binding.
FT   DOMAIN   1   829   MYOSIN HEAD-LIKE.
FT   DOMAIN   830   859   IQ
FT   DOMAIN   886   2017   COILED COIL (POTENTIAL).
FT   NP_BIND   225   232   ATP.
FT   DOMAIN   250   260   25 KDA/50 KDA JUNCTION.
FT   DOMAIN   682   694   50 KDA/20 KDA JUNCTION.
FT   DOMAIN   705   727   ACTIN-BINDING.
FT   DOMAIN   742   758   REACTIVE SULFHYDRYL/ACTIN-BINDING.
FT   DOMAIN   1303   2017   LIGHT MEROMYOSIN (LMM).
FT   DOMAIN   1303   1970   ALPHA-HELICAL TAILPIECE (LMM).
FT   DOMAIN   1971   2017   GLOBULAR TAILPIECE.
FT   VARSPLIC   1   45   MISSING (IN SHORT ISOFORM).
SQ   SEQUENCE   2017 AA; 232016 MW; 73E3CB02BA8F2528 CRC64;

Query Match   9.08; Score 162.5; DB 1; Length 2017;
Best Local Similarity   23.8%; Pred. No. 1.9;
Matches 96; Conservative 68; Mismatches 139; Indels 101; Gaps 18;

Qy   26 LTEISKKITESNAVV------LAVKEVE--TLLASIDELATKAIGKKIG-------- 66
            :::: ||: || :|| | | ||||
Db 1108 VADLKEQLNERRVQVDEMAQLAKREEELTQTLRIDESATKATPAQKAQRELESQOLAEI 1167

Qy   67 NNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTRAKQCSTEPTNKLKSEHA 126
            |||| |:|:| : |: || |:|
Db 1168 QEDLEAEKAARAK------AEKVRDLSEELEALKNELLDSLDTTAAQ--QELRSKREQELA 1221

Qy 127 VIGLDLNLTDDNAQRAILKKHAN------------KDKGAAELEKLFKAVENLSKA----- 172
            ::||:| :|| : | |||| ||
Db 1222 TL----------KKSLEEETVNHHGGVLADMRHKHSQELNSINDQLENLRKAKTVLEKAKG 1271

Qy 173 TLKNAVKELTSPIVHGNNSRKDGNASTNSADESVKGPNJTEISKKITESNAVVLAVK--E 230
Db 1272 TLEAAENADLATELRSVNSSRQENDR----------------RRKQAESQIAELQVKLAE 1314

Qy 231 VETLLASIDELATKRAIGKKIGN--NGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEEL 288
Db 1315 IERARSELQEKCTK-LQQEAENITNQLEEAELKASAAVKSASNMESQLTEAQOLLE-EET 1372

RESULT   9
MLP1_YEAST
ID   MLP1_YEAST   STANDARD;   PRT;   1875 AA.
AC   Q02455;
DT   01-OCT-1993 (Rel. 27, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   MYOSIN-LIKE PROTEIN MLP1.
GN   MLP1 OR YKR095W OR YKR415.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288C;
RX   MEDLINE=93247549; PubMed=8483450;
RA   Koelling R., Nguyen T., Chen E.Y., Botstein D.;
RT   "A new yeast gene with a myosin-like heptad repeat structure.";
RL   Mol. Gen. Genet. 237:359-369(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94205265; PubMed=8154186;
RA   Bou G., Esteban P.F., Baladron V., Gonzalez G.A., Cantalejo J.G.,
RA   Remacha M., Jimenez A., del Rey F., Ballesta J.P.G., Revuelta J.L.;
RT   "The complete sequence of a 15,820 bp segment of Saccharomyces
RT   cerevisiae chromosome XI contains the UB12 and MPL1 genes and three
RT   new open reading frames.";
RL   Yeast 9:1349-1354(1993).
CC   -!- FUNCTION: MYOSIN-LIKE PROTEIN THAT IS PROBABLY INVOLVED IN DNA
CC   REPAIR.
CC   -!- SIMILARITY: SOME, TO THE TPR ONCOGENE.
CC   -!- CAUTION: REF. 2 MISQUOTES THE GENE NAME AS "MPL1".
CC   ---

DR   EMBL; L01992; AAA34783.1; -.
DR   EMBL; X73541; CAA51948.1; -.
DR   EMBL; Z28320; CAA82274.1; -.
DR   PIR; S38173; S38173.
KW   Coiled coil; DNA repair.
FT   DOMAIN   69   487   COILED COIL (POTENTIAL).
FT   DOMAIN   531   1678   COILED COIL (POTENTIAL).
FT   DOMAIN   1834   1866   COILED COIL (POTENTIAL).
FT   CONFLICT   301   301   R -> A (IN REF. 1).
SQ   SEQUENCE   1875 AA; 683A0D34C9066867 CRC64;

Qy 289 KEKIDTRAKQCSTEFTNKLKSEHAVLGLDNLTLDDNAQRAI--------LKKHANKD 335
            :|: || : |:|| || |: :| |:| ||:|
Db 1373 RQKLGLSSK-----LRQIESEKEALQEQLEEDDEAKRNYERKLAEVTTQMQEIKKGAEED 1427

Qy 336 KG-AAELE----KLFKAVENLSK------AAQDTLKNAVKELTS 368
         || |||| |: |:| :|: :||:| :|
Db 1428 ADLAKELEEGKKRLNKDIEALERQVKELIAQNDRLDKSKKKIQS 1471

Query Match   8.9%; Score 159.5; DB 1; Length 1875;
Best Local Similarity   21.3%; Pred. No. 2.4;
Matches 96; Conservative 80; Mismatches 155; Indels 119; Gaps 18;

Qy   26 LTEISKKITESNAVVLAVKEVETLLASIDELATFKAIGKKIGNNGLEANQSKNTSLLSGAY 85
            :::||:|||:::|| |:::| ||||||:|||:|||||
Db 647 ISQITRESTEN--MSLLNKEIQDLYDSKSDISIK-LGKEKSSRILAEERFK---LLSNTL 700

```
RESULT 10
ID    YNP9_CAEEL     STANDARD;      PRT;     705 AA.
AC    P34562;
DT    01-FEB-1994 (Rel. 28, Created)
DT    30-MAY-2000 (Rel. 39, Last sequence update)
DT    30-MAY-2000 (Rel. 39, Last annotation update)
DE    HYPOTHETICAL 80.6 KDA PROTEIN T05G5.9 IN CHROMOSOME III.
GN    T05G5.9.
OS    Caenorhabditis elegans.
OC    Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC    Rhabditidae; Peloderinae; Caenorhabditis.
OX    NCBI_TaxID=6239;
RN    [1]
RP    SEQUENCE FROM N.A.
RC    STRAIN=BRISTOL N2;
RX    MEDLINE=94150718; PubMed=7906398;
RA    Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA    Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA    Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
RA    Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA    Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,
RA    Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA    Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
RA    Sims M., Smaldon N., Smith A., Smith M., Sonnhammer E., Staden R.,
RA    Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA    Waterson R., Watson A., Weinstock L., Wilkinson-Sproat J.,
RA    Wohldman P.;
RT    "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT    elegans.";
RL    Nature 368:32-38(1994).
RN    [2]
RP    REVISIONS.
RA    Durbin R.;
RL    Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
CC    -----------------------------------------------------------------------
CC    This SWISS-PROT entry is copyright. It is produced through a collaboration
CC    between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC    the European Bioinformatics Institute. There are no restrictions on its
CC    use by non-profit institutions as long as its content is in no way
CC    modified and this statement is not removed. Usage by and for commercial
CC    entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC    or send an email to license@isb-sib.ch).
CC    -----------------------------------------------------------------------
DR    EMBL; Z27079; CAA81596.1; -.
DR    PIR; S41009; S41009.
DR    WormPep; T05G5.9; CE21153.
DR    InterPro; IPR000237; GRIP.
DR    Pfam; PF01465; GRIP; 1.
KW    Hypothetical protein; Coiled coil.
FT    DOMAIN    75    137    COILED COIL (POTENTIAL).
FT    DOMAIN   160    509    COILED COIL (POTENTIAL).
FT    DOMAIN   562    641    COILED COIL (POTENTIAL).
SQ    SEQUENCE   705 AA;  80637 MW;  9E89873F5FC04956 CRC64;

Query Match         8.8%;   Score 158.5;  DB 1;  Length 705;
Best Local Similarity   25.0%;   Pred. No. 0.91;
Matches   91;  Conservative  52;  Mismatches  140;  Indels  81;  Gaps  14;

Qy     86 ----AISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNKLKSEHAVLGLDNLTDDNAQRA 141
             ::|:  |||
Db    701 DLTKAENDQLRKPDVLQNTILKQDSKTHETLNEYVSCKSKLSIVETELLNLKEEQKLRV 760

Qy    142 ILKKHANKDKGAAELEK----LFKAVENLSKAAQDTLKNAVKELTSPIVHGNNSRKDGN 196
             |||                 ::|
Db    761 HLEKNLKQELNKLSPEKDSLRIMVTQLQTLQKEREDLLEETRKSCQKKIDELEDALSELK 820

Qy    197 ASTNSADESVKGPNLTEISKKTTE--SNAVVLAVKEVETLLASIDELAT------KA 245
             |:|:  |:|                  ::|                        |:
Db    821 KETSQKDHHIK--QLEEDNSNIEWYQNKIEALKKDYESVITSVVDSKQTDIEKLQYKVS 878

Qy    246 IGKK-----------------------IGNNGLEANQSKNTSLLSGAYA--------- 271
             |||                       ::|
Db    879 LEKEIEEDKIRLHTYNVMDETINDDSLRKELEKSKINLTDAVSQIKEYKDLYETTSOSLQ 938

Qy    272 -------------ISDL-------IAEKLNVLK------NEEL------KEKIDTA 295
                         :|:|         ||||||         |||         |||
Db    939 OTNSKLDESPKDFTNQIKNLTDEKTSLEDKISLLKEQMFNLNNELDLQKKGMEKEKADFK 998

Qy    296 KQCS-TEFTNK-----LKSEHAVLGLDNLTDDNAQRAI-----------LKKHANKDGA 338
             |::|  |          |::|:||                         |||:
Db    999 KRISLQNNNKEVEAVKSEYE-SKLSKIQNDLDQQTIYANTAQNNYEQELQKHADVSKTI 1057

Qy    339 AELEK----LFKAVENLSKAAQDTLKNAVKE 365
             ::|:|       |:|:||
Db   1058 SELREQLHTYKGQVKTLNLSRDQLENALKE 1087
```

```
RESULT 11
ID    RBP2_PLAVB     STANDARD;      PRT;     1251 AA.
AC    Q00799;
DT    01-APR-1993 (Rel. 25, Created)
DT    01-APR-1993 (Rel. 25, Last sequence update)
DT    01-OCT-1996 (Rel. 34, Last annotation update)
DE    RETICULOCYTE BINDING PROTEIN 2 (FRAGMENT).
GN    RBP2.
OS    Plasmodium vivax (strain Belem).
OC    Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX    NCBI_TaxID=31277;
RN    [1]
RP    SEQUENCE FROM N.A.
RX    MEDLINE=93215338; PubMed=1617731;
RA    Galinski M.R., Medina C.C., Ingravallo P., Barnwell J.W.;
RT    "A reticulocyte-binding protein complex of Plasmodium vivax
RT    merozoites.";
RL    Cell 69:1213-1226(1992).
CC    -!- FUNCTION: INVOLVED IN RETICULOCYTE ADHESION. SPECIFICALLY BINDS TO
CC        HUMAN RETICULOCYTE CELLS.
CC    -!- SUBCELLULAR LOCATION: MEMBRANE-BOUND (PROBABLE).
CC    -----------------------------------------------------------------------
CC    This SWISS-PROT entry is copyright. It is produced through a collaboration
CC    between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC    the European Bioinformatics Institute. There are no restrictions on its
CC    use by non-profit institutions as long as its content is in no way
CC    modified and this statement is not removed. Usage by and for commercial

Qy    25 NLFEISKKITESNAVVLAVKEVETLLASIDELATKAIGKKIGNNGLEANQSKNTSLLSGA 84
             |:|:|  ||:||||  |||||||||||||||||||||:||||||||||||||||||||||
Db    94 NOTALMEKLKRAKMSELEQVRKDAENLKLINEKLTTESAKVENNPTEC-----TECLSKS 148

Qy    85 YAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNKLKSEHAVLGLDNLTDDNAQRAILK 144
             |:|:|  :|:|||  ||                          |||||||||||||||||
Db   149 GALIELEKEVF----------------------EWKEK----------------ATRA-- 168

Qy   145 KHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKE--LTSPIVHGNNSRKDGNASTNS 201
             ||:                 ::|                    |||||:||:
Db   169 -----DMISELRDLESKVDQLNRALRDKTEALIKAQEVITENDLEVNNMKKEKN-NTKS 222

Qy   202 ADESVKGPNLTEISKKITESNAVVLAVKEVETLLAS---IDELATKAIGKKIGNNGLEA 257
             :|:|                                  ||||||||||||||||||||
Db   223 SIEKLTEEN-TRLTKALQDEK---IKSADFEARLRSAECRIVELSDQOGNEKIGLARKMA 278

Qy   258 NQSKKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNKLKSEHAVLGLDN 317
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   279 ESENRGRILEEAVDVLKSENEKL-LAKNEFSAKLVSSEK--EFAEFKKKSHFVLEKKG 334

Qy   318 LTDDNAQRAILKKHANKDKGAAELEKL------FKAVENLSKA-----AQDTLKNAV 363
             |||||                        ||||||||          ||||||||
Db   335 KQEDEFRKAIEKLEKSK-VITELEQOADQTRQEHFKTVEDLASSRDKAERLEKTLKVLK 393

Qy   364 KELT 367
             |||
Db   394 SELT 397
```

```
CC   entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC
DR   EMBL; M88098; AAA29744.1; -.
KW   Malaria; Receptor; Membrane.
FT   NON_TER      1      1
FT   NON_TER   1251   1251
SQ   SEQUENCE   1251 AA;  143741 MW;  54BA51C7404AC572 CRC64;
```

```
Qy   Query Match             8.8%;  Score 157.5;  DB 1;  Length 1251;
     Best Local Similarity  20.6%;  Pred. No. 1.9;
     Matches  106;  Conservative  77;  Mismatches  151;  Indels  181;  Gaps  21;

Qy    9  DGNASANSADESVK---GPNLIEISKKITESNAVVLAV------KEVETLLASIDELAT   58
          .:||     |      ||:|| |  ||  ||:|:  |         ||::| |:  ||
Db  671  EGHGNVKQLVENIKELVDENNLSDILKQATGKNEEIQKITHSTLKNKAKTILGHVDTSA-  729

Qy   59  KAIGKKI------------------GNNGLEA-NQSKNTSLLSGAYAISDLI           91
          :|| ||                    | ||| |  :|||||||||||| ||:
Db  730  KYVGIKITPELALTELLGDAKLKTAQELKFESKNNVVLETENMSKNTNELDVHKNIQDAY  789

Qy   92  AEKLNVLKNEEELKEKIDFAKQCST--------------EFTNKLKS-             123
          :|  |:  || ||| |:  | ||               ||:||:|:
Db  790  KVALEILAH--SDEIDIKQKDSSKLIEMGNQIYLKVVLINQYKNKISSIKSKEEAVSVK  846

Qy  124  ------EHAVLG----------LDNLFTDDNAQRAILKKHANKDKGAAELE          157
                :|||:|            |:||| ||:|| |||||||:|| |||:|
Db  847  IGNVSKKHSELSKITCSDKSYDNIIIALEKQTELQNLRNSFTQE---KTNTNSD---SKIE 900

Qy  158  KLFKAVENLSKAAQDTLKNAVKELTSPIVHGNNSRKDGNASTNSADESVKGPNLTEISKK 217
          |:  .|:||| | |   ||| |:| |     :|| |  | |    :: |||| | ||:|
Db  901  KI-------KIDFESLKNALKTLEGE-VNALKASSDNHEHVQSKSEPV-NPALSEIEKE  950

Qy  218  ITESNAVVLAVKEV---ETLLASIDELATKAIGKKIGN------NGLEAN          258
          :||||||||||:|     |||||||||| ||||||| |      |||||
Db  951  ETIDISLNTALDELLKGRTCEVSRYKLIKDZVTKEISDDTELINTIEKNVKAYLAYIKK 1010

Qy  259  ---------QSKNTSLLSGAYAISDLIAEKL------NV                     282
                   | |||||||||||| ||| | :|       |:
Db 1011  NYEDTVQDVLTLNEHENTKQVSNHEPTNFDKSNKSSEELIKAVTDSKTIISKLKGVIIEV 1070

Qy  283  LKNEELEKIDTPAKQCSTEFTNKLKSEHAVLGLDNLTDDNAQRAILKKHANKDGAAELE  342
          |:|||||||||| ||||||||||||||| ||||||||||||||||||||||| ||||||
Db 1071  NENTEMWTIESSAKEIEALY-NELKNKKTSLNEIYQTSNEVKLQEMKSNADK--YIDVS 1126

Qy  343  KLFKAVENLSKA---------AQDTLKNAVKEL                           366
          |:||||||||            |||||||||||
Db 1127  KIFNTVLDTQKSNIVINQHSINNVKDKIKGKLQEL 1161
```

RESULT 12
LMG1_DROME
ID   LMG1_DROME   STANDARD;   PRT;   1639 AA.
AC   P15215; Q24373; Q9V718;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   LAMININ GAMMA-1 CHAIN PRECURSOR (LAMININ B2 CHAIN).
GN   LANB2 OR LAMC1 OR CG3322.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CANTON-S, AND OREGON-R;
RX   MEDLINE=91299161; PubMed=1840513;
RA   Chi H.-C., Juminaga D., Wang S.Y., Hui C.-F.;
RT   "Structure of the Drosophila gene for the laminin B2 chain.";
RL   DNA Cell Biol. 10:451-466(1991).
RN   [2]

RP   SEQUENCE FROM N.A.
RC   STRAIN=OREGON-R;
RX   MEDLINE=89109164; PubMed=2912972;
RA   Chi H.-C., Hui C.-F.;
RT   "Primary structure of the Drosophila laminin B2 chain and comparison
RT   with human, mouse, and Drosophila laminin B1 and B2 chains.";
RL   J. Biol. Chem. 264:1543-1550(1989).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=OREGON-R;
RX   MEDLINE=90037237; PubMed=2808533;
RA   Montell D.J., Goodman C.S.;
RT   "Drosophila laminin: sequence of B2 subunit and expression of all
RT   three subunits during embryogenesis.";
RL   J. Cell Biol. 109:2441-2453(1989).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li J.Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murray D., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [5]
RP   SEQUENCE OF 344-1639 FROM N.A.
RC   STRAIN=OREGON-R;
RX   MEDLINE=88303364; PubMed=3405777;
RA   Chi H.-C., Hui C.-F.;
RT   "cDNA and amino acid sequences of Drosophila laminin B2 chain.";
RL   Nucleic Acids Res. 16:7205-7215(1988).
-!- FUNCTION: BINDING TO CELLS VIA A HIGH AFFINITY RECEPTOR, LAMININ
    IS THOUGHT TO MEDIATE THE ATTACHMENT, MIGRATION, & ORGANIZATION OF
    CELLS INTO TISSUES DURING EMBRYONIC DEVELOPMENT BY INTERACTING
    WITH OTHER EXTRACELLULAR MATRIX COMPONENTS.
-!- SUBUNIT: LAMININ IS A COMPLEX GLYCOPROTEIN, CONSISTING OF THREE
    DIFFERENT POLYPEPTIDE CHAINS (ALPHA, BETA, GAMMA), WHICH ARE BOUND
    TO EACH OTHER BY DISULFIDE BONDS INTO A CROSS-SHAPED MOLECULE
    COMPRISING ONE LONG & THREE SHORT ARMS WITH GLOBULES AT EACH END.
-!- SUBCELLULAR LOCATION: BASEMENT MEMBRANES (MAJOR COMPONENT).
-!- DOMAIN: THE ALPHA-HELICAL DOMAINS I AND II ARE THOUGHT TO INTERACT

```
CC        WITH OTHER LAMININ CHAINS TO FORM A COILED COIL STRUCTURE.
CC   -!- DOMAIN: DOMAINS VI AND IV ARE GLOBULAR.
CC   -!- SIMILARITY: CONTAINS 1 LAMININ N-TERMINAL DOMAIN (DOMAIN VI).
CC   -!- SIMILARITY: CONTAINS 11 LAMININ EGF-LIKE DOMAINS.
CC   -!- SIMILARITY: CONTAINS 1 LAMININ DOMAIN IV.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; M58417; AAA28665.1; -.
DR   EMBL; M25063; AAA28664.1; -.
DR   EMBL; AE003551; AAF50238.1; -.
DR   EMBL; X07806; CAA30665.1; -.
DR   PIR; A31483; MMFFB2.
DR   HSSP; P02468; 1TLE.
DR   FlyBase; FBgn0002528; LanB2.
DR   InterPro; IPR000561; EGF-like.
DR   InterPro; IPR001886; LamNT.
DR   InterPro; IPR000034; Laminin_B.
DR   InterPro; IPR002049; Laminin_EGF.
DR   Pfam; PF00052; laminin_B; 1.
DR   Pfam; PF00053; laminin_EGF; 10.
DR   Pfam; PF00055; laminin_Nterm; 1.
DR   ProDom; PD002082; LamNT; 1.
DR   ProDom; PD003031; Laminin_B; 1.
DR   SMART; SM00180; EGF_Lam; 10.
DR   SMART; SM00001; EGF_like; 1.
DR   SMART; SM00281; LamB; 1.
DR   SMART; SM00136; LamNT; 1.
DR   PROSITE; PS00022; EGF_1; 8.
DR   PROSITE; PS00186; EGF_2; 11.
DR   PROSITE; PS01248; LAMININ_TYPE_EGF; 11.
KW   Laminin EGF-like domain; Basement membrane; Extracellular matrix; Coiled coil;
KW   Glycoprotein; Cell adhesion; Repeat; Signal.
FT   SIGNAL        1     33
FT   CHAIN        34   1639       LAMININ GAMMA-1 CHAIN.
FT   DOMAIN       34    298       LAMININ N-TERMINAL (DOMAIN VI).
FT   DOMAIN      299    523       4.5 X LAMININ EGF-LIKE REPEATS (DOMAIN
FT                               V).
FT   DOMAIN      299    358       LAMININ EGF-LIKE 1.
FT   DOMAIN      359    413       LAMININ EGF-LIKE 2.
FT   DOMAIN      414    460       LAMININ EGF-LIKE 3.
FT   DOMAIN      461    513       LAMININ EGF-LIKE 4.
FT   DOMAIN      514    523       LAMININ EGF-LIKE 5 (N-TERMINAL).
FT   DOMAIN      524    709       LAMININ DOMAIN IV.
FT   DOMAIN      710   1049       6.5 X LAMININ EGF-LIKE REPEATS (DOMAIN
FT                               III).
FT   DOMAIN      710    743       LAMININ EGF-LIKE 5 (C-TERMINAL).
FT   DOMAIN      744    792       LAMININ EGF-LIKE 6.
FT   DOMAIN      793    846       LAMININ EGF-LIKE 7.
FT   DOMAIN      847    901       LAMININ EGF-LIKE 8.
FT   DOMAIN      902    955       LAMININ EGF-LIKE 9.
FT   DOMAIN      956   1003       LAMININ EGF-LIKE 10.
FT   DOMAIN     1004   1049       LAMININ EGF-LIKE 11.
FT   DOMAIN     1050   1609       DOMAIN II AND I.
FT   DOMAIN     1087   1109       COILED COIL (POTENTIAL).
FT   DOMAIN     1144   1247       COILED COIL (POTENTIAL).
FT   DOMAIN     1306   1627       COILED COIL (POTENTIAL).
FT   DISULFID    299    308       BY SIMILARITY.
FT   DISULFID    301    322       BY SIMILARITY.
FT   DISULFID    324    333       BY SIMILARITY.
FT   DISULFID    336    356       BY SIMILARITY.
FT   DISULFID    359    368       BY SIMILARITY.
FT   DISULFID    361    384       BY SIMILARITY.
FT   DISULFID    387    396       BY SIMILARITY.
FT   DISULFID    399    411       BY SIMILARITY.
FT   DISULFID    414    426       BY SIMILARITY.
FT   DISULFID    416    432       BY SIMILARITY.
FT   DISULFID    434    443       BY SIMILARITY.
FT   DISULFID    446    458       BY SIMILARITY.
FT   DISULFID    461    475       BY SIMILARITY.
FT   DISULFID    463    482       BY SIMILARITY.
FT   DISULFID    484    493       BY SIMILARITY.
FT   DISULFID    496    511       BY SIMILARITY.
FT   DISULFID    744    753       BY SIMILARITY.
FT   DISULFID    746    760       BY SIMILARITY.
FT   DISULFID    762    771       BY SIMILARITY.
FT   DISULFID    774    790       BY SIMILARITY.
FT   DISULFID    793    801       BY SIMILARITY.
FT   DISULFID    795    811       BY SIMILARITY.
FT   DISULFID    814    823       BY SIMILARITY.
FT   DISULFID    826    844       BY SIMILARITY.
FT   DISULFID    847    861       BY SIMILARITY.
FT   DISULFID    849    868       BY SIMILARITY.
FT   DISULFID    871    880       BY SIMILARITY.
FT   DISULFID    883    899       BY SIMILARITY.
FT   DISULFID    902    919       BY SIMILARITY.
FT   DISULFID    904    926       BY SIMILARITY.
FT   DISULFID    928    937       BY SIMILARITY.
FT   DISULFID    940    953       BY SIMILARITY.
FT   DISULFID    956    968       BY SIMILARITY.
FT   DISULFID    958    975       BY SIMILARITY.
FT   DISULFID    977    986       BY SIMILARITY.
FT   DISULFID    989   1001       BY SIMILARITY.
FT   DISULFID   1004   1016       BY SIMILARITY.
FT   DISULFID   1006   1022       BY SIMILARITY.
FT   DISULFID   1024   1033       BY SIMILARITY.
FT   DISULFID   1036   1047       BY SIMILARITY.
FT   DISULFID   1050   1050       INTERCHAIN (PROBABLE).
FT   DISULFID   1053   1053       INTERCHAIN (PROBABLE).
FT   DISULFID   1631   1631       INTERCHAIN (PROBABLE).
FT   CARBOHYD    147    147       N-LINKED (GLCNAC...) (POTENTIAL).
```

Query Match   8.7%;   Score 157;   DB 1;   Length 1639;
Best Local Similarity 24.1%;   Pred. No. 2.7;
Matches 104;   Conservative 61;   Mismatches 158;   Indels 108;   Gaps 21;

```
Qy      2  ACNNSGKDGNASANSAD------ESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDE  55
Db   1107  ARDNSGDGGQQTYAEVIDDLHKHLDSVR-EHLVSADKFQADANGEIDRARQNYTILDQTE 1165

Qy     56  LATKAIGKKI-------GNNGLEANQSKNTSLLSGAYAISDL----IAEKLNV---L  98
Db   1166  NAKKELQQALDLLNDEGGAQALARAKEKSVEFGQQOSEQISDISREARALADKLESEAQFDL 1225

Qy     99  KN-EELKEKIDTAKQCS-------TEFTNKLKSEHAVLGLDNLTDDNAQORAILRKHANKD 150
Db   1226  KNAKDAKDAVERAHQLAKSAIDLQLKIGTELRSE---VGLE------LSHVKQS 1270

Qy    151  KGA---AELEKLFKAVENLSKRAAQDTLKNAVKELTSPIVHGNNSRKDGNASTNSADESVK 207
Db   1271  LGTVVQOTSKEALRKANEVYDTAL--TLLNDVRQTOPEIDISQLKKDAVAANERADELLK 1328

Qy    208  GPNLTEISKKITESNAVVLAVKEVETLLASIDELATRAIGKKIIGNNGLEANQSKNTSLLS 267
Db   1329  -QITELS----NSNGELFADPETEQEL------TEALLKR-----AEQOQLEDIELLE 1370

Qy    268  GAYAISDLIAEKL-----NVLK-----------NEELKEKIDTAKQCSTEFTN 304
Db   1371  RAKAAAHDKATKRAVEBOQGDNTLKEANNTYEKLAGPOSDVQORSSESAEKALQTVPNIEKEION 1430

Qy    305  K----LKSEHAVLGLD---NLTDDNAQORAILK--KIHANKDKGAAAELEKLFKAVENLSKAA 355
Db   1431  AESLISQAEEALDGANNNANEAKKNAQEAQLKYAEQASKD-----AELIRRKANETKVA 1484

Qy    356  QDTLKNAVKEL 366
Db   1485  ARNLREEADQL 1495
```

```
ID   A60D_DROME     STANDARD;      PRT;   1013 AA.
AC   P91927; Q9W160;
DT   30-MAY-2000 (Rel. 39, Created)
DT   20-AUG-2001 (Rel. 40, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   CALCIUM-BINDING MITOCHONDRIAL PROTEIN ANON-60DA.
GN   ANON-60DA OR CG4589.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C., Smith H.O.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [2]
RP   SEQUENCE OF 626-944 FROM N.A.
RC   TISSUE=Ovary;
RX   MEDLINE=99168769; PubMed=10071211;
RA   Caggese C., Ragone G., Perrini B., Moschetti R., de Pinto V.,
RA   Caizzi R., Barsanti P.;
RT   "Identification of nuclear genes encoding mitochondrial proteins:
RT   isolation of a collection of D. melanogaster cDNAs homologous to
RT   sequences in the Human Gene Index database.";
RL   Mol. Gen. Genet. 261:64-70(1999).
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL.
CC   -!- SIMILARITY: CONTAINS 2 EF-HAND CALCIUM-BINDING DOMAINS.
CC   -!- CAUTION: REF.2 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO
CC       FRAMESHIFTS IN POSITIONS 920 AND 930.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
```

```
DE   MYOSIN TYPE II HEAVY CHAIN 1.
GN   MYO2 OR SPC645.05C.
OS   Schizosaccharomyces pombe (Fission yeast).
OC   Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC   Schizosaccharomycetales; Schizosaccharomycetaceae;
OC   Schizosaccharomyces;
OX   NCBI_TaxID=4896;
RN   [1]
RP   SEQUENCE FROM N.A., AND FUNCTION.
RC   STRAIN=972;
RX   MEDLINE=98075862; PubMed=9415380;
RA   May K.M., Watts F.Z., Jones N., Hyams J.S.;
RT   "Type II myosin involved in cytokinesis in the fission yeast,
RT   Schizosaccharomyces pombe.";
RL   Cell Motil. Cytoskeleton 38:385-396(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=972;
RA   Wood V., Rajandream M.A., Barrell B.G., Rieger M.;
RL   Submitted (MAR-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: REQUIRED FOR CELL DIVISION. IT IS A COMPONENT OF THE
CC       CDC12 'SPOT', A STRUCTURE THOUGHT TO MARK THE SITE OF SEPTATION.
CC       MAY WORK IN CONJUNCTION WITH MYO3.
CC   -!- SUBUNIT: BINDS TO CDC4 AND RLC1
CC   -!- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 IQ DOMAIN.
CC       -----------------------------------------------------------------------
CC       This SWISS-PROT entry is copyright. It is produced through a collaboration
CC       between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC       the European Bioinformatics Institute. There are no restrictions on its
CC       use by non-profit institutions as long as its content is in no way
CC       modified and this statement is not removed. Usage by and for commercial
CC       entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC       or send an email to license@isb-sib.ch).
CC       -----------------------------------------------------------------------
DR   EMBL; U75357; AAC49908.1; -.
DR   EMBL; AL049498; CAB39901.1; -.
DR   HSSP; P08799; 1MND.
DR   InterPro; IPR000048; myosin_head.
DR   InterPro; IPR001609; myosin_head.
DR   Pfam; PF00063; myosin_head; 1.
DR   Pfam; PF00612; IQ; 1.
DR   PRINTS; PR00193; MYOSINHEAVY.
DR   ProDom; PD000355; myosin_head; 1.
DR   SMART; SM00015; IQ; 1.
DR   SMART; SM00242; MYSc; 1.
DR   PROSITE; PS00016; IQ; 1.
KW   Myosin; Actin-binding; ATP-binding; Coiled coil; Calmodulin-binding;
KW   Alkylation.
FT   DOMAIN          1    757       MYOSIN HEAD-LIKE.
FT   DOMAIN        758    787       IQ.
FT   DOMAIN        875   1244       COILED COIL (POTENTIAL).
FT   NP_BIND       170    177       ATP (POTENTIAL).
FT   DOMAIN        634    656       ACTIN-BINDING (BY SIMILARITY).
FT   DOMAIN        734    748       ACTIN-BINDING (BY SIMILARITY).
FT   MOD_RES       674    674       ALKYLATION (BY SIMILARITY).
FT   CONFLICT     1337   1337       S -> R (IN REF. 1).
SQ   SEQUENCE   1526 AA;  176430 MW;  D71D51D6578192BA CRC64;

Query Match             8.6%;
Best Local Similarity  23.3%;
Matches   88;  Conservative  76;  Mismatches  139;  Indels  74;  Gaps  18;

              Score 154;  DB 1;  Length 1526;
              Pred. No. 3.3;

Qy      8 KDGNASANSADESVKGPN-LTEISKKITESNAVVLAVKEV-----ETLLASIDE------ 55
Db    842 KQQNSKSEVERDLVETNNSLTAVENLLTTERAIALDKEEILRRTQERLANIEDSFSEIKQ 901

Qy     56 ----LATKAIGKKIGNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTPAK 111
Db    902 QNENLQRESASLKQINNELESELLEKTS-----KVETLLSEQ------NELKEKLSLEE 949

Qy    112 QCSTEFTNKLKS--EHAVIGLDNLTDDNAQ-RAILKKHANKDKGAAELEKLFRAVENLSK 168
```

```
Db    950 KDLLDLTKGELESLRENNATVLSEKAEFNEQCKSLQEFIVTKD---AELDKLTKYISDYKT 1006
Qy    169 AAQDTLKNAVKELTSPIVHGNNSRKDGNASTNSADESVKGPNLTEISKKITESNAVVLA- 227
Db   1007 EIQE------MRLTNQKMNEKSIQQEGSLS-----ESLKR------VKKLERENSTLISD 1049
Qy    228 -----VKEVETLLASIDELATKAIGKKIGNNGLEANQSKNTSLLSGAYAISDL-IAEKL 280
Db   1050 VSILKQQKEELSVLKGVQOELTINNLEEKV--NYLEADVKQLPKLKKELESLNDKDQLYQL 1107
Qy    281 NVLKNEELKEKIDTAKQCSTEFTNKLKS----EHAVIGLDNLTDDNAQRAILKK-HANK 334
Db   1108 QATKNRELEAKV---KEC---LNNIKSLTKELENKEEKCQNLSDASLKYIELQEIHENL 1160
Qy    335 DKGAAELEKLFKAVENL 351
Db   1161 LLKVSDLENYKKYEGL 1177

RESULT 15
CEAK_ECOLI
ID   CEAK_ECOLI     STANDARD;      PRT;    548 AA.
AC   Q47502; P75615;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   COLICIN K.
GN   CKA.
OS   Escherichia coli.
OG   Plasmid ColK-K235.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Izard J., Chartier M., Baty D.;
RL   Submitted (MAY-1995) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96074330; PubMed=7592493;
RA   Pilsl H., Braun V.;
RT   "Strong function-related homology between the pore-forming colicins K
RT   and 5.";
RL   J. Bacteriol. 177:6973-6977(1995).
CC   -!- FUNCTION: THIS COLICIN IS A CHANNEL-FORMING COLICIN. THIS CLASS OF
CC       TRANSMEMBRANE TOXINS DEPOLARIZE THE CYTOPLASMIC MEMBRANE, LEADING
CC       TO DISSIPATION OF CELLULAR ENERGY.
CC   -!- FUNCTION: COLICINS ARE POLYPEPTIDE TOXINS PRODUCED BY AND ACTIVE
CC       AGAINST ESCHERICHIA COLI AND CLOSELY RELATED BACTERIA.
CC   -!- SIMILARITY: BELONGS TO THE CHANNEL FORMING COLICIN FAMILY.
CC       -----------------------------------------------------------------------
CC       This SWISS-PROT entry is copyright. It is produced through a collaboration
CC       between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC       the European Bioinformatics Institute. There are no restrictions on its
CC       use by non-profit institutions as long as its content is in no way
CC       modified and this statement is not removed. Usage by and for commercial
CC       entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC       or send an email to license@isb-sib.ch).
CC       -----------------------------------------------------------------------
DR   EMBL; X87834; CAA61099.1; -.
DR   EMBL; U27452; AAB41288.1; -.
DR   InterPro; IPR000293; Channel_colicin.
DR   Pfam; PF01024; Colicin; 1.
DR   PRINTS; PR00280; CHANNELCOLICIN.
DR   ProDom; PD002657; Channel_colicin; 1.
DR   PROSITE; PS00276; CHANNEL_COLICIN; 1.
KW   Antibiotic; Bacteriocin; Plasmid; Transmembrane.
FT   TRANSMEM      505    525       POTENTIAL.
FT   CONFLICT      155    155       S -> G (IN REF. 2).
SQ   SEQUENCE    548 AA;  59661 MW;  2E676B9D088CCF86 CRC64;
```

Query Match   8.5%;   Score 153;   DB 1;   Length 548;
Best Local Similarity   20.8%;   Pred. No. 1.2;
Matches   90; Conservative   78; Mismatches 161; Indels 104; Gaps 19;

```
Qy     5 NSGKDGNASANSADESVKGPNLTEISK---------------------------KITESNAVL------  41
         |:| ||::|| :|      :|                                    :: ::
Db    46 NNGGQGNSNSGGSTSTVMKTGESYLTPWGDVVINNDGLPVMNGIVWTEENSTLVDNPFGG 105

Qy    42 AVKEVETLLASIDELATKAIGKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNE 101
         |: :|::: ||:| :|||| ||:|| |     |  ||::||    :| || :
Db   106 VSRVLNSLISDMPSLFAESSGNN-NNTASVNTAPTNAQVSDMDKSSKVVSNVIN--EKQ 162

Qy   102 ELKEKIDT----AKQCSTEFTNKLKSEHAVLGLDNLTDD----------------NAQR 140
         :|||| |     ||                                         |:|
Db   163 KQKNKIATQISEKQKKIEEMKKVFKHHSYHGITDLERDVDELQKKSNQLDADISKLNSYK 222

Qy   141 AILKK---HANKDKGAAELEKLFKAVENLSKAAQDTL---KNAVKELTSPIVHGNNSRKD 194
         || || | ||| |  ||| | | :| | |||    |   | ||
Db   223 NTLQSKIGDVNKQKEAEE------KARENAEVAEHEILNEBKQAVAEAEKRLA------E 270

Qy   195 GNASTNSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELAT----KAIGKK 249
         |:|    |: | |||    ||:||:    :|:  ||| | | |||||:      |||||
Db   271 AKAELAKAESDVQ--------SKQATVSR---VAGELENAQKSVDVKVTGFPGWRDVQKK 319

Qy   250 IGNNGLEANQSKNTSL---LSGAYAISD---------LIAEKLNVLKNEELKEKIDTA 295
         :| |: ::  ::                         ||:|||:|| ||:||| :::
Db   320 L-QRQLEAKQAEYSAVENELKNAVSFRDGKAAEVKEAEQKLKEAQDALEKSQIKDAVDTM 378

Qy   296 KQCSTEFTNKLKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAV--ENLSK 353
         |: :||              |:|          |   :| | | | ||: |     | ||
Db   379 VGFYQYITEQYGEKYA----KIAQDLAEKSKGKKIQGVDEALAAFEK-YKNVLDKKFSK 432

Qy   354 AAQDTLKNAVKEL 366
         :| |||||::::|
Db   433 VDRDAIFNALESV 445
```

Search completed: March 18, 2002, 10:11:55
Job time: 973 sec

OM protein - protein search, using sw model

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

Run on:　　March 18, 2002, 10:10:54 ; Search time 124.19 Seconds
(without alignments)
433.434 Million cell updates/sec

Title:　　　1797
Perfect score:
Sequence:　1 MACNNSGKDGNASANSADES............ENLSKAAQDTLKNAVKELTS 368

Scoring table:　BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched:　473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters:　473505
Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing:　Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :　SPTREMBL_17:*

1:	sp_archea:*	
2:	sp_bacteria:*	
3:	sp_fungi:*	
4:	sp_human:*	
5:	sp_invertebrate:*	
6:	sp_mammal:*	
7:	sp_mhc:*	
8:	sp_organelle:*	
9:	sp_phage:*	
10:	sp_plant:*	
11:	sp_rodent:*	
12:	sp_virus:*	
13:	sp_vertebrate:*	
14:	sp_unclassified:*	

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	% DB	DB ID	Description
1	901	192	2	Q9RR52	Q9rr52 borrelia bu
2	901	209	2	Q4717	Q4717 borrelia bu
3	899	191	2	Q93P1	Q93p1 borrelia bu
4	889	192	2	031117	031117 borrelia bu
5	663	194	2	031122	031122 borrelia bu
6	661	193	2	P94234	P94234 borrelia bu
7	653.5	212	2	Q44705	Q44705 borrelia bu
8	646	202	2	008231	008231 borrelia ta
9	646	211	2	Q44720	Q44720 borrelia bu
10	644	201	2	Q9RQR8	Q9rqr8 borrelia bu
11	642	193	2	P94237	P94237 borrelia bu
12	639	191	2	Q44726	Q44726 borrelia bu
13	623.5	192	2	Q9R7B1	Q9r7b1 borrelia bu
14	623.5	193	2	Q9RR53	Q9rr53 borrelia bu
15	623	185	2	031123	031123 borrelia bu
16	621.5	192	2	Q93P3	Q93p3 borrelia bu
17	620	200	2	Q9XDH4	Q9xdh4 borrelia bu
18	619	193	2	P94236	P94236 borrelia bu
19	616	209	2	Q44883	Q44883 borrelia bu
20	614	191	2	Q9S3P0	Q9s3p0 borrelia bu
21	612	188	2	Q9XDH3	Q9xdh3 borrelia bu
21	610.5	203	2	050620	050620 borrelia af
22	609.5	200	2	Q9RQR9	Q9rqr9 borrelia bu
23	609.5	205	2	P96505	P96505 borrelia af
24	607.5	201	2	031115	031115 borrelia bu
25	607.5	193	2	P96571	P96571 borrelia ja
26	607.5	201	2	Q44670	Q44670 borrelia ja
27	607.5	212	2	P96572	P96572 borrelia af
28	605.5	191	2	P70818	P70818 borrelia bu
29	605.5	194	2	P94229	P94229 borrelia ja
30	605.5	201	2	P96573	P96573 borrelia ja
31	605	178	2	044979	044979 borrelia bu
32	605	178	2	044995	044995 borrelia bu
33	603.5	210	2	044719	044719 borrelia bu
33	602.5	201	2	P96516	P96516 borrelia ja
34	602.5	201	2	Q9S3P2	Q9s3p2 borrelia ja
35	601.5	192	2	P96572	P96572 borrelia bu
35	601.5	191	2	P70819	P70819 borrelia bu
36	601.5	201	2	Q9S504	Q9s504 borrelia bu
37	600	201	2	031120	031120 borrelia bu
38	599	184	2	Q9S3P2	Q9s3p2 borrelia ja
39	598.5	191	2	P96572	P96572 borrelia bu
40	596.5	182	2	Q9R7B2	Q9r7b2 borrelia bu
41	596.5	212	2	044718	044718 borrelia ja
42	594.5	201	2	P96514	P96514 borrelia bu
43	594.5	181	2	034124	034124 borrelia ja
44	592	184	2	034120	034120 borrelia bu
45	591	211	2	Q9Z6C7	Q9z6c7 borrelia va

ALIGNMENTS

RESULT 1
Q9RR52

ID　Q9RR52　　PRELIMINARY;　　PRT;　　192 AA.
AC　Q9RR52;
DT　01-MAY-2000 (TrEMBLrel. 13, Created)
DT　01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT　01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE　OUTER SURFACE PROTEIN C (FRAGMENT)
OSPC.
OS　Borrelia burgdorferi (Lyme disease spirochete).
OC　Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX　NCBI_TaxID=139;
RN　[1]
RP　SEQUENCE FROM N.A.
RC　STRAIN=OC5;
RX　MEDLINE=99091544; PubMed=9872945;
RA　Wang I.N., Dykhuizen D.E., Qiu W., Dunn J.J., Bosler E.M., Luft B.J.;
RT　"Genetic diversity of ospC in a local population of Borrelia
RT　burgdorferi sensu stricto.";
RL　Genetics 151:15-30(1999).
DR　EMBL; AF029864; AAB86547.1; -.
DR　InterPro; IPR001800; Lipoprotein_6.
DR　Pfam; PF01441; Lipoprotein_6; 1.
DR　ProDom; PD001149; Lipoprotein_6; 1.
FT　NON_TER　1
FT　NON_TER　192
SQ　SEQUENCE　192 AA;　20555 MW;　FF16D409D58C01C5 CRC64;

Query Match　　50.1%;　　Score 901; DB 2; Length 192;
Best Local Similarity　98.9%;　　Pred. No. 3.6e-35;
Matches 184;　Conservative 2;　Mismatches 0;　Indels 0;　Gaps 0;

Qy　1　MACNNSGKDGNASANSADESVKGPNLTEISKKITTESNAVVLAVKEVETLLASIDELATKA 60
　　　　::||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　7　ISCNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKA 66

Qy　61　IGKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNK 120
　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　67　IGKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNK 126

Qy　121　LKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKE 180

```
Db  127 LKSEHAVIGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKE 186
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy  181 LTSPIV 186
        ||||||
Db  187 LTSPIV 192
```

RESULT 2

```
ID   Q44717
AC   Q44717;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C (FRAGMENT).
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=N40;
RX   MEDLINE=94314484; PubMed=8039931;
RA   Stevenson B., Bockenstedt L.K., Barthold S.W.;
RT   "Expression and gene sequence of outer surface protein C of Borrelia
RT   burgdorferi reisolated from chronically infected mice.";
RL   Infect. Immun. 62:3568-3571(1994).
DR   EMBL; U04240; AAC45538.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
SQ   SEQUENCE   209 AA;  22393 MW;  3707A47DAA7736FCA CRC64;
```

PRELIMINARY;   PRT;   209 AA.

```
Query Match    50.1%;  Score 901;  DB 2;  Length 209;
Best Local Similarity   98.9%;  Pred. No. 4e-35;
Matches  184;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    1 MACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKA 60
        ::||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   17 ISCNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKA 76

Qy   61 IGKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNK 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   77 IGKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNK 136

Qy  121 LKSEHAVIGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKE 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  137 LKSEHAVIGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKE 196

Qy  181 LTSPIV 186
        ||||||
Db  197 LTSPIV 202
```

RESULT 3

```
ID   Q9S3P1
AC   Q9S3P1;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C (FRAGMENT).
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=28691;
RX   MEDLINE=96296448; PubMed=8709845;
RA   Livey I., Gibbs C.P., Schuster R., Dorner F.;
RT   "Evidence for lateral transfer and recombination in OspC variation in
RT   Lyme disease Borrelia.";
RL   Mol. Microbiol. 18:257-269(1995).
DR   EMBL; L42894; AAB37002.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
SQ   SEQUENCE   191 AA;  20340 MW;  8CA1A6ACF17AED8F CRC64;
```

PRELIMINARY;   PRT;   191 AA.

```
Query Match    50.0%;  Score 899;  DB 2;  Length 191;
Best Local Similarity   100.0%;  Pred. No. 4e-35;
Matches  184;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    3 CNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKAIG 62
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 CNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKAIG 60

Qy   63 KKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNKLK 122
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 KKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNKLK 120

Qy  123 SEHAVIGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKELT 182
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 SEHAVIGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKELT 180

Qy  183 SPIV 186
        ||||
Db  181 SPIV 184
```

RESULT 4

```
ID   O31117
AC   O31117;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C (FRAGMENT).
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=OC7;
RA   Wang I.-N., Dykhuizen D.E., Dunn J.J., Luft B.J.;
RL   Submitted (OCT-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF029866; AAB86549.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
SQ   SEQUENCE   192 AA;  20684 MW;  1F0C6BB66E291F6B CRC64;
```

PRELIMINARY;   PRT;   192 AA.

```
Query Match    49.5%;  Score 889;  DB 2;  Length 192;
Best Local Similarity   97.8%;  Pred. No. 1.3e-34;
Matches  184;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy    1 MACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKA 60
        :: |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7 ISCNNSRDGNASTNSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKA 66

Qy   61 IGKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNK 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   67 IGKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNK 126

Qy  121 LKSEHAVIGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKE 180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  127 LKSEHAVIGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKE 186
```

```
Qy   181 LTSPIV 186
         ||||||
Db   187 LTSPIV 192


RESULT 5
ID   O31122       PRELIMINARY;      PRT;   194 AA.
AC   O31122;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C (FRAGMENT).
OS   OSPC.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   STRAIN=OC12;
RA   Wang I.-N., Dykhuizen D.E., Dunn J.J., Luft B.J.;
RL   Submitted (OCT-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF029871; AAB86554.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
FT   NON_TER        1      1
FT   NON_TER      194    194
SQ   SEQUENCE   194 AA;  20640 MW;  00A5E6E2D2CE0F7F CRC64;

Query Match             36.9%;  Score 663;  DB 2;  Length 194;
Best Local Similarity   73.9%;  Pred. No. 3.5e-24;
Matches   139;  Conservative   22;  Mismatches   25;  Indels    2;  Gaps    2;

Qy    1 MACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKA 60
        :||||||||||||||||| ||:|||||||||||||||||||||||:|||||||||||||||
Db    7 ISCNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDELATKA 66

Qy   61 IGKKI-GNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKN-EELKEKIDTAKQCSTEFT 118
        |||||  ||:||||  ||||||||||||||||||||||||  ||||||||||||||||||
Db   67 IGKKIQQNGGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKEKIENAKKCSEDFT 126

Qy  119 NKLKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAV 178
        |:|| ||| |:::||::|:||  :|||  |||||||||||||||||||::|||:  ::|
Db  127 KKLEGEHAQLGIENVDENAKKAILITDAAKDKGAAELEKLFKAVENLAKAAKEMLANSV 186

Qy  179 KELTSPIV 186
        ||||||||
Db  187 KELTSPIV 194


RESULT 6
ID   P94234       PRELIMINARY;      PRT;   193 AA.
AC   P94234;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C (FRAGMENT).
OS   OSPC.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=28354;
RX   MEDLINE=96296448; PubMed=8709845;
RA   Livey I., Gibbs C.P., Schuster R., Dorner F.;
RT   "Evidence for lateral transfer and recombination in OspC variation in
RT   Lyme disease Borrelia.";
RL   Mol. Microbiol. 18:257-269(1995).
DR   EMBL; L42895; AAB37003.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
FT   NON_TER        1      1
FT   NON_TER      193    193
SQ   SEQUENCE   193 AA;  20411 MW;  05B68720F061E2A0 CRC64;

Query Match             36.8%;  Score 661;  DB 2;  Length 193;
Best Local Similarity   74.7%;  Pred. No. 4.3e-24;
Matches   139;  Conservative   20;  Mismatches   25;  Indels    2;  Gaps    2;

Qy    3 CNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKAIG 62
        :|||||||||| ||:||||||||||||||||||||||||||||:|||||||||||||||||
Db    1 CNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDELATKAIG 60

Qy   63 KKI-GNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKN-EELKEKIDTAKQCSTEFTNK 120
        ||| |||  ||:||||  |||||||||||||||||||||  ||||||||||||||||||||
Db   61 KKIQQNGGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKEKIENAKKCSEDFTKK 120

Qy  121 LKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKE 180
        |: |||| |:::||::|:||  :|||  ||||||||||||||||||||::|||:  ::|||
Db  121 LEGEHAQLGIENVDENAKKAILITDAAKDKGAAELEKLFKAVENLAKAAKEMLANSVKE 180

Qy  181 LTSPIV 186
        ||||||
Db  181 LTSPIV 186


RESULT 7
ID   Q44705       PRELIMINARY;      PRT;   212 AA.
AC   Q44705;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C.
OS   OSPC.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=2591;
RX   MEDLINE=94041630; PubMed=8225587;
RA   Padula S.J., Sampieri A., Dias F., Szczepanski A., Ryan R.W.;
RT   "Molecular characterization and expression of p23 (OspC) from a North
RT   American strain of Borrelia burgdorferi.";
RL   Infect. Immun. 61:5097-5105(1993).
DR   EMBL; U01892; AAA16057.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
SQ   SEQUENCE   212 AA;  22270 MW;  FB2EF2673A384276 CRC64;

Query Match             36.4%;  Score 653.5;  DB 2;  Length 212;
Best Local Similarity   72.4%;  Pred. No. 1.1e-23;
Matches   139;  Conservative   21;  Mismatches   29;  Indels    3;  Gaps    2;

Qy    1 MACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKA 60
        :||||||||||||||||| ||:|||||||||||||||||||||||:|||||||||:|:||
Db   17 ISCNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDEVAKKA 76

Qy   61 IGKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKN-EELKEKIDTAKQCSTEFTN 119
        ||  | ||| || || |||||||||:||:||||||||||:  |||||||||||| ::|
Db   77 IGNLIAQNGLNAGANQNGSLLAGAYVISTLLAEKLDGLKNSEELKEIEDAKKCNKAFTD 136

Qy  120 KLKSEHAVLGLDN--LTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNA 177
        |||| |||||:||    ::|||::|||:|  :|:|||:|||:|||  :|||||||:||:||
Db  137 KLKSSHAELGIANGAASDANAKAAILKTNGTKDKGAQELEKLFESVKNLSKAAQEFLNNS 196
```

```
Qy    178 VKELTSPIYHGN 189
          |||||:|| |
Db    197 VKELTSPVVAEN 208

RESULT 8
ID   O08231      PRELIMINARY;    PRT;   202 AA.
AC   O08231;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C (FRAGMENT).
OS   Borrelia tanukii.
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=56146;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Fi81T;
RA   Masuzawa T., Komikado T., Fukui T., Yanagihara Y.;
RT   "B. tanukii(Strain Fi81t) ospC, partial cds.";
RL   Submitted (JAN-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB000354; BAA19087.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
NON_TER      1      1
NON_TER    202    202
SEQUENCE   202 AA;  21536 MW;  F309F684C6896IE7 CRC64;

Query Match             35.9%;  Score 646;  DB 2;  Length 202;
Best Local Similarity   70.8%;  Pred. No. 2.2e-23;
Matches  133;  Conservative  22;  Mismatches  31;  Indels   2;  Gaps   2;

Qy     1 MACNNSGKDG-NASANSADESVKGPNLTEISRKITESNAVVLAVKEVETLLASIDELATK 59
         :||||||| | :|| :||||||||||||||||: ||||||||||||| ||||||||||
Db    11 ISCNNSGKGGDSASTNPVDESAKGPNLTEISKKITDSNAIVLAVKEVETLISSIDELANK 70

Qy    60 AIGKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKN-EELKEKIDTAKQCSTEFT 118
         ||||| |||| ||| | |:||||||||:||||:||||||  |||||||||||:|:||
Db    71 AIGKKIQNNGLDTLSDKNASLLSGAYVISTLITEKLNGLKNSEDLKEKIEEAKKCSAAFT 130

Qy   119 NKLKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAV 178
          ||||:|:|: ||  ||||||||||||||||||||||||||||||||||||||:|:||
Db   131 NKLKNSHAQLGIEAATDDDAKKAILKTHVTKDKGAAEFEEKLFKSVEILIKAAQEALANSI 190

Qy   179 KELTSPIV 186
          |||||:|
Db   191 KELTSPVV 198

RESULT 9
ID   Q44720      PRELIMINARY;    PRT;   211 AA.
AC   Q44720;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OSPC.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=25015;
RX   MEDLINE=95154673; PubMed=7851744;
RA   Stevenson B., Barthold S.W.;
RT   "Expression and sequence of outer surface protein C among North
RT   American isolates of Borrelia burgdorferi.";
RL   FEMS Microbiol. Lett. 124:367-372(1994).
DR   EMBL; U04282; AAC45540.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
SEQUENCE   211 AA;  22603 MW;  69984BA6D8743ED5 CRC64;

Query Match             35.9%;  Score 646;  DB 2;  Length 211;
Best Local Similarity   72.3%;  Pred. No. 2.3e-23;
Matches  136;  Conservative  18;  Mismatches  32;  Indels   2;  Gaps   2;

Qy     1 MACNNSGKDGN-ASANSADESVKGPNLTEISRKITESNAVVLAVKEVETLLASIDELATK 59
         :|||||||||| || :|||||||||||||:||||: |||||||||||   ||:|||||||
Db    17 ISCNNSGKGDGNAASTNPADESVKGPNLTEISKKITDSNAVVLAVKEVGALLTSIDELATK 76

Qy    60 AIGKKI-GNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFT 118
         ||||| :|||| ||| | |:||||||| | |:|| |  :|||||||:|:||
Db    77 AIGKKIHQNNGLDTENNIINGSLLAGAYATAISTLIITQKLGGLKNEELEKIAAVKKCSEEFT 136

Qy   119 NKLKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAV 178
          ||||:|: | | |  ||||||||: |  ||||||:| ||||||||||||||:|:||
Db   137 NKLKSSHFELGKQDAQDDAKKAILRTHNTKDKGAEELDKLFKAVENLSKAAKEMLSNSV 196

Qy   179 KELTSPIV 186
          |||||:|
Db   197 KELTSPVV 204

RESULT 10
ID   Q9RQR8      PRELIMINARY;    PRT;   201 AA.
AC   Q9RQR8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C PRECURSOR (FRAGMENT).
OS   Borrelia burgdorferi (Lyme disease spirochete).
OG   Plasmid cp26.
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=3B6';
RX   MEDLINE=20002545; PubMed=10531219;
RA   Hofmeister E.K., Glass G.E., Childs J.E., Persing D.H.;
RT   "Population dynamics of a naturally occurring heterogeneous mixture of
RT   Borrelia burgdorferi clones.";
RL   Infect. Immun. 67:5709-5716(1999).
DR   EMBL; AF074465; AAD23912.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
KW   Plasmid.
NON_TER    201    201
SEQUENCE   201 AA;  21457 MW;  806F19829510IB07 CRC64;

Query Match             35.8%;  Score 644;  DB 2;  Length 201;
Best Local Similarity   73.4%;  Pred. No. 2.7e-23;
Matches  135;  Conservative  22;  Mismatches  25;  Indels   2;  Gaps   2;

Qy     1 MACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKA 60
         :||||||| |::||||||||||||||||||||||||||||||||:|:||||||||||||||
Db    17 ISCNNSGKGDGNTSANSADESVKGPNLTEISKKITDSNAVVLAVKEITLLASIDELATKA 76

Qy    61 IGKKI-GNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKN-EELKEKIDTAKQCSTEFT 118
         |||| :|||| |:   ||:|||:||| |  :|| :||  :|||||||:|  ||||:|||
Db    77 IGKKIQQNGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKEKIENAKKCSEDFT 136

Qy   119 NKLKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAV 178
          ||| |||:| | |  |::||:||| |  ||||||||||||||||||||:||||:|:||
Db   137 KKLEGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLAKAAKEMLANSV 196

Qy   179 KELTSPIV 186
          |||||:|
Db   197 KELTSPVV 204
```

```
Qy   179 KELT 182
         |||
Db   197 KELT 200
```

RESULT 11

ID   P94237
AC   P94237;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C (FRAGMENT).
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=25015;
RX   MEDLINE=96296448; PubMed=8709845;
RA   Livey I., Gibbs C.P., Schuster R., Dorner F.;
RT   "Evidence for lateral transfer and recombination in OspC variation in
RT   Lyme disease Borrelia.";
RL   Mol. Microbiol. 18:257-269(1995).
DR   EMBL; L42898; AAB37007.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
FT   NON_TER   1   1
FT   NON_TER   193   193
SQ   SEQUENCE   193 AA; 20677 MW; C9500D959E13590D CRC64;

PRELIMINARY;   PRT;   193 AA.

Query Match   35.7%; Score 642; DB 2; Length 193;
Best Local Similarity   72.6%; Pred. No. 3.2e-23;
Matches 135; Conservative 16; Mismatches 33; Indels   2; Gaps   2;

```
Qy     3 CNNSGKDGN-ASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKAI 61
           ||||||||  :|||||| ||||||||||||||||||:|||||||||||||||||||||||
Db     1 CNNSGKDGNAASTNPADESVKGPNLTEISKKITDSNTVLAVKEVEALLTSIDELATKAI 60

Qy    62 GKKI-GNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNK 120
          |||| ||||:||:||||||||||||||||||  :||| ||:||||||||||||||||||||
Db    61 GKKIHQNNGLDTENNHNGSLLAGAYAISTLITQKIGGLKNEELKEKIAAVKKCSEEFTNK 120

Qy   121 LKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKE 180
          ||||   | :| |::||:||: | |||||||||||:||||||||||||||:||:| ||||
Db   121 LKSSHELGKQDADDDAKKAILRTHNTKDKGAEELDKLFKPVENLSKAAKEMLSNSVKE 180

Qy   181 LTSPIV 186
          |||:|
Db   181 LTSPVV 186
```

●

RESULT 12

ID   O44726
AC   O44726;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C (FRAGMENT).
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=297;
RX   MEDLINE=94314437; PubMed=8039891;
RA   Fung B.P., McHugh G.L., Leong J.M., Steere A.C.;
RT   "Humoral immune response to outer surface protein C of Borrelia
RT   burgdorferi in Lyme disease: role of the immunoglobulin M response in

---

RT   the serodiagnosis of early infection.";
RL   Infect. Immun. 62:3213-3221(1994).
DR   EMBL; U08284; AAA21460.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
FT   NON_TER   1   1
SQ   SEQUENCE   191 AA; 20150 MW; C49A4030F0A28717 CRC64;

Query Match   35.6%; Score 639; DB 2; Length 191;
Best Local Similarity   73.9%; Pred. No. 4.4e-23;
Matches 136; Conservative 20; Mismatches 26; Indels   2; Gaps   2;

```
Qy     5 NSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKAI 64
          |||| || ||||||| |||||||||||||||:|||||||||||||||||||||||||
Db     1 NSGKGGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVEIELLASIDELATKAIGKK 60

Qy    65 I-GNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKN-EELKEKIDTAKQCSTEFTNKLK 122
          | :|||||:||||||||||||||||||  :||| |  ::||||||:|||||||:||:| ||
Db    61 IQQNGLAVEAGHNGTLLAGAYTISKLLITQKLDGLKNSEKLKEKIENAKKCGSEDFTKKLE 120

Qy   123 SEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKELT 182
          |||||||||:||||:||: | |||||||||||:|||||||||||:|| ||::|:|||||
Db   121 GEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLAKAAKEMLANSVKELT 180

Qy   183 SPIV 186
          ||||
Db   181 SPIV 184
```

RESULT 13

ID   Q9R7B1
AC   Q9R7B1;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C (FRAGMENT).
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TETS;
RX   MEDLINE=97478003; PubMed=9336916;
RA   Ras N.M., Postic D., Foretz M., Baranton G.;
RT   "Borrelia burgdorferi sensu stricto, a bacterial species 'made in the
RU.S.A.'?";
RL   Int. J. Syst. Bacteriol. 47:1112-1117(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TETS;
RA   Marti-Ras N., Postic D., Foretz M., Baranton G.;
RL   Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U91798; AAB81895.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
FT   NON_TER   1   1
SQ   SEQUENCE   192 AA; 20297 MW; 67705D2A20AAF764 CRC64;

PRELIMINARY;   PRT;   192 AA.

Query Match   34.7%; Score 623.5; DB 2; Length 192;
Best Local Similarity   69.5%; Pred. No. 2.3e-22;
Matches 130; Conservative 24; Mismatches 32; Indels   1; Gaps   1;

```
Qy     1 MACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKA 60
          :::|||||||| ||||||||| |||||||||||||||:|||||||| ||:|||||||||||
Db     3 ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLAVKEVEALLSSIDEIAAKA 62

Qy    61 IGKKI-GNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTN 119
```

```
RESULT 14

ID   Q9RR53          PRELIMINARY;      PRT;    193 AA.
AC   Q9RR53;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C (FRAGMENT).
GN   OSPC.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=OC1;
RX   MEDLINE=99091544; PubMed=9872945;
RA   Wang I.N., Dykhuizen D.E., Qiu W., Dunn J.J., Bosler E.M., Luft B.J.;
RT   "Genetic diversity of ospC in a local population of Borrelia
RT   burgdorferi sensu stricto.";
RL   Genetics 151:15-30(1999).
DR   EMBL; AF029860; AAB85543.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
FT   NON_TER       1      1
FT   NON_TER     193    193
SQ   SEQUENCE   193 AA;  20502 MW;  5EFDB5AFF8986D1E CRC64;

Query Match             34.7%;  Score 623.5;  DB 2;  Length 193;
Best Local Similarity   69.5%;  Pred. No. 2.3e-22;
Matches  130;  Conservative   24;  Mismatches   32;  Indels    1;  Gaps    1;

Qy          1 MACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKA 60
              ::|||||||| |||||||||| ||||||||||||||||||||||||||:|||||||||||
Db          7 ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA 66

Qy         61 IGKKI-GNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTN 119
              |||||  ||||||| |||||| :||  ||:|||:|||| ||:||||||||||||:||||||
Db         67 IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSEFTFTN 126

Qy        120 KLKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAAELEKLFKAVENLSKAAQDTLKNAVK 179
              |||:|||||||| ||||||||| ||||| ||:|| ||:|| ||||||||:|||||:|| :|
Db        123 KLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 182

Qy        180 ELTSPIV 186
              |||||:|
Db        183 ELTSPVV 189


RESULT 15

ID   O31123          PRELIMINARY;      PRT;    185 AA.
AC   O31123;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C (FRAGMENT).
GN   OSPC.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=QC13;
RA   Wang I.-N., Dykhuizen D.E., Dunn J.J., Luft B.J.;
RL   Submitted (OCT-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF029872; AAB86555.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
FT   NON_TER       1      1
FT   NON_TER     185    185
SQ   SEQUENCE   185 AA;  19673 MW;  58D6FEE3C7769CAF CRC64;

Query Match             34.7%;  Score 623;  DB 2;  Length 185;
Best Local Similarity   72.1%;  Pred. No. 2.3e-22;
Matches  129;  Conservative   24;  Mismatches   24;  Indels    2;  Gaps    2;

Qy          1 MACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKA 60
              ::|||||||| ||||||||||||||||||||||||||:|||||||||||:|||||||||||
Db          7 ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA 66

Qy         61 IGKKI-GNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKN-EELKEKIDTAKQCSTEFT 118
              |||||  ||||||| |||||| :||  ||:|||:|||| ||  |:||||||||||:||||
Db          7 IGKKIQQNNGLGLAVEAGHNGTLLAGAYTISKLITQELDGLKNSEKLKEKIENAKKCSEDFT 126

Qy        119 NKLKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAAELEKLFKAVENLSKAAQDTLKNA 177
              |:||:|||||||| |||||||||| |||| ||:|| ||:|| ||||||||:|||||:||
Db        127 KKLEGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLAKAAKEFLNNS 185

Qy        178 VKELTSPIV 186
              ||||||:|
Db        187 VKELTSPVV 193
```

Search completed: March 18, 2002, 10:10:55

Job time: 978 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5

Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 09:54:33 :

Title: US-09-596-746A-42
Perfect score: 1798
Sequence: 1 MACNNSGKDGNASANSADES..............AVENLAKAAKEMLANSVKEL 368

Scoring table: BLOSUM62
Searched: Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Post-processing: Minimum Match 0%
Maximum DB seq length: 0
Minimum DB seq length: 2000000000
          Listing first 45 summaries

Database :   A_Geneseq_1101:*
          1: /SIDS8/gcgdata/geneseq/geneseqp/AA1980.DAT:*
          2: /SIDS8/gcgdata/geneseq/geneseqp/AA1981.DAT:*
          3: /SIDS8/gcgdata/geneseq/geneseqp/AA1982.DAT:*
          4: /SIDS8/gcgdata/geneseq/geneseqp/AA1983.DAT:*
          5: /SIDS8/gcgdata/geneseq/geneseqp/AA1984.DAT:*
          6: /SIDS8/gcgdata/geneseq/geneseqp/AA1985.DAT:*
          7: /SIDS8/gcgdata/geneseq/geneseqp/AA1986.DAT:*
          8: /SIDS8/gcgdata/geneseq/geneseqp/AA1987.DAT:*
          9: /SIDS8/gcgdata/geneseq/geneseqp/AA1988.DAT:*
          10: /SIDS8/gcgdata/geneseq/geneseqp/AA1989.DAT:*
          11: /SIDS8/gcgdata/geneseq/geneseqp/AA1990.DAT:*
          12: /SIDS8/gcgdata/geneseq/geneseqp/AA1991.DAT:*
          13: /SIDS8/gcgdata/geneseq/geneseqp/AA1992.DAT:*
          14: /SIDS8/gcgdata/geneseq/geneseqp/AA1993.DAT:*
          15: /SIDS8/gcgdata/geneseq/geneseqp/AA1994.DAT:*
          16: /SIDS8/gcgdata/geneseq/geneseqp/AA1995.DAT:*
          17: /SIDS8/gcgdata/geneseq/geneseqp/AA1996.DAT:*
          18: /SIDS8/gcgdata/geneseq/geneseqp/AA1997.DAT:*
          19: /SIDS8/gcgdata/geneseq/geneseqp/AA1998.DAT:*
          20: /SIDS8/gcgdata/geneseq/geneseqp/AA1999.DAT:*
          21: /SIDS8/gcgdata/geneseq/geneseqp/AA2000.DAT:*
          22: /SIDS8/gcgdata/geneseq/geneseqp/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB ID	Description
1	1793	99.7	367	22	AAB62719	Borrelia sp chimer
2	1790	99.6	391	22	AAB62735	Borrelia sp chimer
3	1557.5	86.6	368	22	AAB62718	Borrelia sp chimer
4	1554.5	86.5	392	22	AAB62734	Borrelia sp chimer
5	1538	85.5	368	22	AAB62717	Borrelia sp chimer
6	1530	85.1	391	22	AAB62736	Borrelia sp chimer
7	1485.5	82.6	377	22	AAB62713	Borrelia sp chimer
8	1484	82.5	373	22	AAB62711	Borrelia sp chimer
9	1481	82.4	397	22	AAB62729	Borrelia sp chimer
10	1480	82.3	369	22	AAB62716	Borrelia sp chimer
11	1479.5	82.3	400	22	AAB62739	Borrelia sp chimer
12	1477	82.1	393	22	AAB62731	Borrelia sp chimer
13	1250	69.5	378	22	AAB62712	Borrelia sp chimer
14	1248.5	69.4	374	22	AAB62710	Borrelia sp chimer
15	1245.5	69.3	398	22	AAB62728	Borrelia sp chimer
16	1244.5	69.2	370	22	AAB62715	Borrelia sp chimer
17	1244	69.2	401	22	AAB62738	Borrelia sp chimer
18	1241.5	69.0	394	22	AAB62730	Borrelia sp chimer
19	1220	67.9	369	22	AAB62714	Borrelia sp chimer
20	1217	67.7	393	22	AAB62732	Borrelia sp chimer
21	1167.5	64.9	560	22	AAB62726	Borrelia sp chimer
22	1159.5	64.5	393	22	AAB62724	Borrelia sp chimer
23	1156.5	64.3	408	22	AAB62737	Borrelia sp chimer
24	1144	63.6	384	22	AAB62725	Borrelia sp chimer
25	1136	63.2	401	22	AAB62733	Borrelia sp chimer
26	1101.5	61.3	410	22	AAB62740	Borrelia sp chimer
27	1100.5	61.2	386	22	AAB62727	Borrelia sp chimer
28	901	50.1	184	22	AAB62705	B burgdorferi ospC
29	899	50.0	184	22	AAB62702	B burgdorferi ospC
30	899	50.0	191	15	AAR60884	B burgdorferi 28691 Osp
31	889	49.4	191	22	AAB62706	B burgdorferi ospC
32	876	48.7	193	22	AAB62709	B burgdorferi ospC
33	699	38.9	193	22	AAB62709	B burgdorferi ospC
34	699	38.8	587	16	AAR75746	B31 Osp-A/antigen
35	698.5	38.8	587	16	AAR75739	B31 Osp-A/antigen
36	665	37.0	466	16	AAR75740	B31 outer surface
37	657.5	36.6	212	16	AAW11934	Outer surface prot
38	651.5	36.2	194	15	AAR60885	B burgdorferi OspC
39	644.5	35.8	192	22	AAB62706	B burgdorferi ospC
40	644.5	35.8	210	16	AAW11935	Outer surface prot
41	640.5	35.6	190	22	AAB62707	B burgdorferi ospC
42	640	35.6	193	15	AAR60888	Borrelia ZS7 antig
43	640	35.6	211	18	AAW41022	B. burgdorferi sen
44	639.5	35.5	192	15	AAR60889	Borrelia 297 antig
45	639.5	35.6	210	16	AAR75727	B. burgdorferi str

ALIGNMENTS

RESULT 1
AAB62719

AAB62719 AAB62719 standard; Protein; 367 AA.
XX
ID
AC   AAB62719;
XX
DT   03-APR-2001 (first entry)
XX
DE   Borrelia sp chimeric ospC protein SEQ ID NO: 42.
XX
KW   Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
XX
PN   WO200078566-A1.
XX
PD   28-DEC-2000.
XX
PF   19-JUN-2000; 2000WO-US16915.
XX
PR   18-JUN-1999;   99US-0140042.
XX
PA   (UYN) UNIV NEW YORK STATE RES FOUND.
PA   (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI   Dattwyler RJ, Seinost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
DR   WPI; 2001-050013/06.
DR   N-PSDB; AAF29023.
XX
PT   Compositions of OspC polypeptides from strains of Borrelia which cause
PT   Lyme disease are used to immunize animals and detect immune responses
PT   to Lyme disease -

RESULT 2

```
AAB62735;
AAB62735 standard; Protein; 391 AA.

AC   AAB62735;
XX
DT   03-APR-2001  (first entry)
XX
DE   Borrelia sp chimeric ospC protein SEQ ID NO: 74.
XX
KW   Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
XX
PN   WO200078966-A1.
XX
PD   28-DEC-2000.
XX
PF   19-JUN-2000; 2000WO-US16915.
XX
PR   18-JUN-1999; 99US-0140042.
XX
PA   (UNYN ) UNIV NEW YORK STATE RES FOUND.
PA   (BROO-) BROOK BIOTECHNOLOGIES INC.
PI   Dattwyler RJ, Seinost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
PS   Claim 43; Page 102-103; 160pp; English.
XX
CC   The present invention provides compositions comprising ospC proteins and
CC   chimeric ospC proteins from members of the Borrelia genus. These may be
CC   Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC   vaccines against Borrelia infection, which is spread by ticks and leads
CC   to Lyme disease.
XX
SQ   Sequence   367 AA;

Query Match             99.7%;  Score 1793;  DB 22;  Length 367;
Best Local Similarity   100.0%;
Matches  367;  Conservative    0;  Mismatches   0;  Indels    0;  Gaps    0;
                                   Pred. No. 1.2e-108;

Qy          2 ACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKAI 61
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 acnnsgkdgnasansadesvkgpnlteiskkitesnavvlavkevetllasidelatkai 60

Qy         62 GKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNKL 121
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 gkkignngleanqskntsllsgayaisdliaeklnvlkneelkekidtakqcsteftnkl 120

Qy        122 KSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKEL 181
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ksehavlgldnltddnaqrailkkhankdkgaaeleklfkavenlskaaqdtlknavkel 180

Qy        182 TSPIVHGNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDEL 241
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 tspivhgnnsgkdgntsansadesvkgpnlteiskkitesnavvlavkeietllasidel 240

Qy        242 ATKAIGKKIQQNGGLAVEAGHNGTLLAGAYTISKLITQKLIDGLKNSEKLKEKIENAKKCS 301
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 atkaigkkiqqnggilaveaghngtllagaytiskliitqkldglknseklkekienakkcs 300

Qy        302 EDFTKKLEGEHAQLGIENVTDENAKKAILTDAAKDKGAAELEKLFKAVENLAKAAKEML 361
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 edftkklegehaqlgienvtdenakkailtdaakdkgaaeleklfkavenlakaaakeml 360

Qy        362 ANSVKEL 368
              |||||||
Db        361 ansvkel 367
```

RESULT 3

```
AAB62718
AAB62718;
AAB62718 standard; Protein; 368 AA.

ID   AAB62718 standard; Protein; 368 AA.
AC   AAB62718;
XX
DT   03-APR-2001  (first entry)
XX
DE   Borrelia sp chimeric ospC protein SEQ ID NO: 40.
XX
KW   Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
XX
PN   WO200078966-A1.
XX
PD   28-DEC-2000.
XX
PF   19-JUN-2000; 2000WO-US16915.
XX
PR   18-JUN-1999; 99US-0140042.
XX
DR   WPI; 2001-050113/06.
DR   N-PSDB; AAP29039.
XX
PT   Compositions of OspC polypeptides from strains of Borrelia which cause
PT   Lyme disease are used to immunize animals and detect immune responses
PT   to Lyme disease -
XX
SQ   Sequence   391 AA;
XX
PS   Claim 43; Page 142; 160pp; English.
XX
CC   The present invention provides compositions comprising ospC proteins and
CC   chimeric ospC proteins from members of the Borrelia genus. These may be
CC   Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC   vaccines against Borrelia infection, which is spread by ticks and leads
CC   to Lyme disease.
XX

Query Match             99.6%;  Score 1790;  DB 22;  Length 391;
Best Local Similarity   99.7%;            Pred. No. 2e-108;
Matches  366;  Conservative    1;  Mismatches   0;  Indels    0;  Gaps    0;

Qy          2 ACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKAI 61
              :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 scnnsgkdgnasansadesvkgpnlteiskkitesnavvlavkevetllasidelatkai 84

Qy         62 GKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNKL 121
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 gkkignngleanqskntsllsgayaisdliaeklnvlkneelkekidtakqcsteftnkl 144

Qy        122 KSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKEL 181
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 ksehavlgldnltddnaqrailkkhankdkgaaeleklfkavenlskaaqdtlknavkel 204

Qy        182 TSPIVHGNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDEL 241
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 tspivhgnnsgkdgntsansadesvkgpnlteiskkitesnavvlavkeietllasidel 264

Qy        242 ATKAIGKKIQQNGGLAVEAGHNGTLLAGAYTISKLITQKLIDGLKNSEKLKEKIENAKKCS 301
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 atkaigkkiqqnggilaveaghngtllagaytiskliitqkldglknseklkekienakkcs 324

Qy        302 EDFTKKLEGEHAQLGIENVTDENAKKAILTDAAKDKGAAELEKLFKAVENLAKAAKEML 361
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 edftkklegehaqlgienvtdenakkailtdaakdkgaaeleklfkavenlakaaakeml 384

Qy        362 ANSVKEL 368
              |||||||
Db        385 ansvkel 391
```

RESULT 4

```
AAB62734
ID    AAB62734 standard; Protein; 392 AA.
XX
AC    AAB62734;
XX
DT    03-APR-2001  (first entry)
XX
PN    WO200078966-A1.
XX
OS    Chimeric - Borrelia sp.
OS    Chimeric - Borrelia sp.
XX
KW    Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
DE    Borrelia sp chimeric ospc protein SEQ ID NO: 72.
XX
PI    Dattwyler RJ, Selnost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
DR    WPI; 2001-050113/06.
DR    N-PSDB; AAF29022.
XX
PA    (UYNY ) UNIV NEW YORK STATE RES FOUND.
PA    (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
CC    Compositions of OspC polypeptides from strains of Borrelia which cause
CC    Lyme disease are used to immunize animals and detect immune responses
CC    to Lyme disease -
XX
PT    The present invention provides compositions comprising ospc proteins and
PT    chimeric ospC proteins from members of the Borrelia genus. These may be
CC    Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC    vaccines against Borrelia infection, which is spread by ticks and leads
CC    to Lyme disease.
XX
PS    Claim 43; Page 99-100; 160pp; English.
XX
SQ    Sequence  368 AA;
```

Query Match   86.6%; Score 1557.5; DB 22; Length 368;
Best Local Similarity   87.5%; Pred. No. 2.1e-93;
Matches 321; Conservative 20; Mismatches 25; Indels   1; Gaps   1;

```
Qy    2 ACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATRAI 61
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 acnnsgkdgnasansadesvkgpnlteiskkitesnavvlavkevetllasidelatkai 60

Qy   62 GKKIGNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNKL 121
        ||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db   61 gkkigngleangskntsllsgayaisdliaeklnviknee!kekidtakqcsteftnkl 120

Qy  122 KSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKEL 181
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 ksehavlgldnltddnaqrailkkhankdkgaaeleklfkavenlskaaqdtlknavkel 180

Qy  182 TSPIVHGNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDEL 241
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 tspivhgnnsgkdgntsansadesvkgpnlteiskkitesnavvlavkeietllasidel 240

Qy  242 ATKAIGKKIQONGGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKEKIENAKKCS 301
        ||||||||||| ||||||||||||||| || |||:||||||||||||||||||||||||
Db  241 a-kaigkkikndvslgneadhngsllsgayilsnlitkkisaikdsgelkaelekakkcs 299

Qy  302 EDFTKKLEGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLAKAAAKEML 361
        |:|| |||||:|:| || |||| |||| ||||||||||||||||||||||||||||||
Db  300 eeftaklkgehtdlgkegvtddnakkaiilktnndktkgadelekIfesvkniskaakemI 359

Qy  362 ANSVKEL 368
        |||||||
Db  360 tnsvkel 366
```

```
XX
PD    28-DEC-2000.
XX
PF    19-JUN-2000; 2000WO-US16915.
XX
PR    18-JUN-1999;  99US-0140042.
XX
PA    (UYNY ) UNIV NEW YORK STATE RES FOUND.
PA    (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI    Dattwyler RJ, Selnost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
DR    WPI; 2001-050113/06.
DR    N-PSDB; AAF29038.
XX
CC    Compositions of OspC polypeptides from strains of Borrelia which cause
CC    Lyme disease are used to immunize animals and detect immune responses
CC    to Lyme disease -
XX
PT    The present invention provides compositions comprising ospc proteins and
PT    chimeric ospC proteins from members of the Borrelia genus. These may be
CC    Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC    vaccines against Borrelia infection, which is spread by ticks and leads
CC    to Lyme disease.
XX
PS    Claim 43; Page 139-140; 160pp; English.
XX
SQ    Sequence  392 AA;
```

Query Match   86.5%; Score 1554.5; DB 22; Length 392;
Best Local Similarity   87.2%; Pred. No. 3.5e-93;
Matches 320; Conservative 21; Mismatches 25; Indels   1; Gaps   1;

```
Qy    2 ACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATRAI 61
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   25 scnnsgkdgnasansadesvkgpnlteiskkitesnavvlavkevetllasidelatkai 84

Qy   62 GKKIGNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNKL 121
        |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db   85 gkkigngleangskntsllsgayaisdliaeklnviknee!kekidtakqcsteftnkl 144

Qy  122 KSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKEL 181
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  145 ksehavlgldnltddnaqrailkkhankdkgaaeleklfkavenlskaaqdtlknavkel 204

Qy  182 TSPIVHGNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDEL 241
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  205 tspivhgnnsgkdgntsansadesvkgpnlteiskkitesnavvlavkeietllasidel 264

Qy  242 ATKAIGKKIQONGGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKEKIENAKKCS 301
        ||||||||||| ||||||||||||||| || |||:||||||||||||||||||||||||
Db  265 a-kaigkkikndvslgneadhngsllsgayilsnlitkkisaikdsgelkaelekakkcs 323

Qy  302 EDFTKKLEGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLAKAAAKEML 361
        |:|| |||||:|:| || |||| |||| ||||||||||||||||||||||||||||||
Db  324 eeftaklkgehtdlgkegvtddnakkaiilktnndktkgadelekIfesvkniskaakemI 383

Qy  362 ANSVKEL 368
        |||||||
Db  384 tnsvkel 390
```

RESULT 5

```
AAB62717
ID    AAB62717 standard; Protein; 368 AA.
XX
AC    AAB62717;
XX
DT    03-APR-2001  (first entry)
XX
DE    Borrelia sp chimeric ospC protein SEQ ID NO: 38.
XX
```

RESULT 6

AAB62736
ID   AAB62736 standard; Protein; 391 AA.
XX

Left column:

KW   Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
XX
PN   WO200078966-A1.
XX
PD   28-DEC-2000.
XX
PF   19-JUN-2000; 2000WO-US16915.
XX
PR   18-JUN-1999;   99US-0140042.
XX
PI   Dattwyler RJ, Seinost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
PA   (UYNY) UNIV NEW YORK STATE RES FOUND.
PA   (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PR   WPI: 2001-050113/06.
PR   N-PSDB; AAF29021.
XX
PS   Claim 43; Page 97; 160pp; English.
XX
CC   The present invention provides compositions comprising ospC proteins and
CC   chimeric ospC proteins from members of the Borrelia genus. These may be
CC   Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC   vaccines against Borrelia infection, which is spread by ticks and leads
CC   to Lyme disease.
XX
SQ   Sequence   368 AA;

Query Match   85.5%; Score 1538; DB 22; Length 368;
Best Local Similarity   86.7%; Pred. No. 3.8e-92;
Matches 319; Conservative 20; Mismatches 27; Indels   2; Gaps   2;

Qy   1 MACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKA 60
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 macnnsgkdgnasansadesvkgpnlteiskkitesnavvlavkevetllasidelatka 60

Qy   61 IGKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNK 120
             ||||||||||||||||| |||||||||||||||| ||||||||||||||||||||||||
Db   61 igkkignngleanqskntsllsgayaisdliaeklnvlkneelkekidtakqcsteftnk 120

Qy   121 LKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKE 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 lksehavlgldnltddnaqrailkkhankdkgaaeleklfkavenlskaaqdtlknavke 180

Qy   181 LTSPIYHGNSNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDE 240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 ltspiyhgnsnsrkdgnastnsadesvkgpnlteiskkitesnavvlavkevetllaside 240

Qy   241 LATKAIGKKIQQNGGLAVEAGHNGTILLAGAYTISKLITQKLDGLKNSEKLKEKIENAKKC 300
             |:|||| |||| | |||||||||||||||||||||||||||||||||||||||||||||
Db   241 latkaigkki-qnngleangsknts1lsgayaisd1iaeklnvlkn-eelkekidtakqc 298

Qy   301 SEDFTKLEGEHAQLGIENVTDENAKKAILITDAARDKGAAELEKLFKAVENLAKAAAEM 360
             : || |: ||||||||||||:| ||||||||||||||||||||||||||||||||:|
Db   299 steftnklksehavlgldnltddnaqrailkkhankdkgaaeleklfkavenlskaaqdt 358

Qy   361 LANSVKEL 368
             |:|:|||
Db   359 lknavkel 366

Right column:

AC   AAB62736;
XX
DT   03-APR-2001 (first entry)
XX
DE   Borrelia sp chimeric ospC protein SEQ ID NO: 76.
XX
KW   Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
XX
PN   WO200078966-A1.
XX
PD   28-DEC-2000.
XX
PF   19-JUN-2000; 2000WO-US16915.
XX
PR   18-JUN-1999;   99US-0140042.
XX
PI   Dattwyler RJ, Seinost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
PA   (UYNY) UNIV NEW YORK STATE RES FOUND.
PA   (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
DR   WPI: 2001-050113/06.
DR   N-PSDB; AAF29040.
XX
PS   Claim 43; Page 144-145; 160pp; English.
XX
CC   The present invention provides compositions comprising ospC proteins and
CC   chimeric ospC proteins from members of the Borrelia genus. These may be
CC   Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC   vaccines against Borrelia infection, which is spread by ticks and leads
CC   to Lyme disease.
XX
SQ   Sequence   391 AA;

Query Match   85.1%; Score 1530; DB 22; Length 391;
Best Local Similarity   86.4%; Pred. No. 1.4e-91;
Matches 317; Conservative 21; Mismatches 27; Indels   2; Gaps   2;

Qy   2 ACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKAI 61
             : ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   25 scnnsgkdgnasansadesvkgpnlteiskkitesnavvlavkevetllasidelatkai 84

Qy   62 GKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNKL 121
             |||||||||||||||| |||||||||||||| |||||||||||||||||||||||||||
Db   85 gkkignngleanqskntsllsgayaisdliaeklnvlkneelkekidtakqcsteftnkl 144

Qy   122 KSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKEL 181
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   145 ksehavlgldnltddnaqrailkkhankdkgaaeleklfkavenlskaaqdtlknavkel 204

Qy   182 TSPIYHGNSNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDE 241
             ||||||||||| ||||| | ||||||||||||||||||||||||||||||||||||||
Db   205 tspivhgnsnsrkdgnastnsadesvkgpnlteiskkitesnavvlavkevetllaside 264

Qy   242 ATKAIGKKIQQNGGLAVEAGHNGTILLAGAYTISKLITQKLDGLKNSEKLKEKIENAKKCS 301
             |||||||||: ||||||||||||||||| |||:||:| ||:| ||::|: ||||||
Db   265 atkaigkki-qnngleangsknts1lsgayaisd1iaeklnvlkn-eelkekidtakqcs 322

Qy   302 EDFTKLEGEHAQLGIENVTDENAKKAILITDAARDKGAAELEKLFKAVENLAKAAAEMEML 361
             : || |: ||||||||||||:| |||||||||||||||||||||||||||||||::|
Db   323 teftnklksehavlgldnltddnaqrailkkhankdkgaaeleklfkavenlskaaqdtl 382

Qy   362 ANSVKEL 368
             |:|:|||
Db   383 knavkel 389

RESULT   7
AAB62713
XX
ID   AAB62713 standard; protein; 377 AA.
XX
AC   AAB62713;
XX
DT   03-APR-2001 (first entry)
XX
DE   Borrelia sp chimeric ospC protein SEQ ID NO: 30.
XX
KW   Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
XX
XX
PN   WO200078966-A1.
XX
PD   28-DEC-2000.
XX
PF   19-JUN-1999;   2000WO-US16915.
XX
PR   18-JUN-1999;   99US-0140042.
XX
PA   (UYNY) UNIV NEW YORK STATE RES FOUND.
PA   (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI   Dattwyler RJ, Selnost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
DR   WPI; 2001-050113/06.
DR   N-PSDB; AAF29017.
XX
PT   Compositions of OspC polypeptides from strains of Borrelia which cause
PT   Lyme disease are used to immunize animals and detect immune responses
XX
PS   Claim 43; Page 86-87; 160pp; English.
XX
CC   The present invention provides compositions comprising ospC proteins and
CC   chimeric ospC proteins from members of the Borrelia genus. These may be
CC   Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC   vaccines against Borrelia infection, which is spread by ticks and leads
CC   to Lyme disease.
XX
SQ   Sequence   377 AA;

Query Match   82.6%;   Score 1485.5; DB 22;   Length 377;
Best Local Similarity   82.5%;   Pred. No. 9.8e-89;
Matches 311;   Conservative 23;   Mismatches 34;   Indels   9;   Gaps   2;

Qy   1 MACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKA 60
             ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||| ||
Db   1 macnnsgkdgntsansaadesvkgpnlteinkkitdsnavllavkeveallssidelaaka 60

Qy   61 IGKKI-GNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTN 119
             |||| | |||||||||||||||||||||||||| ||||||||||||||||| |||:|| ||
Db   61 igkkihqngngldteennhngsllagayaistliaqkldglknegikekidaakkcsetftn 120

Qy   120 KLKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVK 179
             ||||||||:||||||||||| ::|| :|| :||||||||||||||| |||||||||||||
Db   121 klkekhtdlgkegvtdadakeallkangtktkgaeelgklfesveviskaakemlansvk 180

Qy   180 ELTSPIVHG------NNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEI 231
             |:|||:|:| |||||||||||||||||||||||||||||||||||
Db   181 eltspvvaespkkpsmvnnsgkdgntsansadesvkgpnlteiskkitesnavvlavkei 240

Qy   232 ETLLASIDELATKAIGKKIQONGGLAVEEAGHNGTLLAGAYTISKLITOKLDGLKNSEKLK 291
             |||||||||:||||||| |||||||||||||||||||||||||||||||||||||||||
Db   241 etllasidelatkaigkkiqqngglaveaghngtllagayt iskliqgkldglknsekik 300

Qy   292 EKIENAKKCSEDFTKLEGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVE 351
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 ekienakkcsedftkklegehaqlgienvtdenakkailitdaakdkgaaelekifkave 360

Qy   352 NLAKAAKEMLANSVKEL 368
             ||||||||||||||||||
Db   361 nlakaakemlansvkel 377

RESULT   8
AAB62711
XX
ID   AAB62711 standard; Protein; 373 AA.
XX
AC   AAB62711;
XX
DT   03-APR-2001 (first entry)
XX
DE   Borrelia sp chimeric ospC protein SEQ ID NO: 26.
XX
KW   Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
XX
XX
PN   WO200078966-A1.
XX
PD   28-DEC-2000.
XX
PF   19-JUN-1999;   2000WO-US16915.
XX
PR   18-JUN-1999;   99US-0140042.
XX
PA   (UYNY) UNIV NEW YORK STATE RES FOUND.
PA   (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI   Dattwyler RJ, Selnost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
DR   WPI; 2001-050113/06.
DR   N-PSDB; AAF29015.
XX
PT   Compositions of OspC polypeptides from strains of Borrelia which cause
PT   Lyme disease are used to immunize animals and detect immune responses
XX
PS   Claim 43; Page 81; 160pp; English.
XX
CC   The present invention provides compositions comprising ospC proteins and
CC   chimeric ospC proteins from members of the Borrelia genus. These may be
CC   Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC   vaccines against Borrelia infection, which is spread by ticks and leads
CC   to Lyme disease.
XX
SQ   Sequence   373 AA;

Query Match   82.5%;   Score 1484;   DB 22;   Length 373;
Best Local Similarity   83.4%;   Pred. No. 1.2e-88;
Matches 311;   Conservative 22;   Mismatches 34;   Indels   6;   Gaps   2;

Qy   2 ACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKAI 61
             ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||| ||
Db   1 acnnsgkdgntsansadesvkgpnlteiskkitdsnavllavkeveallssidelaakai 60

Qy   62 GKKI-GNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNK 120
             |||| | |||||||||||||||||||||||| ||||||||||||||||| |||:||||
Db   61 gkkihqngngldteynhngsllagayaistliqkldglknegikekidaakkcsetftnk 120

Qy   121 LKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKE 180
             |||||||:||||||||||| ::||:|| :||||||||||||||| ||||||||||||||
Db   121 lkekhtdlgkegvtdadakeallktngtktkgaeelgklfesveviskaakemlansvke 180

Qy   181 LTSPIVHG-----NNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLL 235
             |:|||:|:| ||||||||||||||||||||||||||||||||||||||
Db   181 ltspvvaespamvnnsgkdgntsansadesvkgpnlteiskkitesnavvlavkeietll 240

Left column

```
Qy  236 ASIDELATKAIGKKIQONGGLAVEAGHNGTLLLAGAYTISKLIITOKLIDGLKNSEKLKEKIE 295
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 asidelatkaikqngglaveaghngtllagaytiskiitqkldglknseklkekie 300

Qy  296 NAKKCSEDFTKLEGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLAK 355
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 nakkcsedftkklegehaqlgienvtdenakkailitdaakdkgaaeleklfkavenlak 360

Qy  356 AAKEMLANSVKEL 368
        |||||||||||||
Db  361 aakemlansvkel 373
```

RESULT 9

```
ID   AAB62729 standard; Protein; 397 AA.
XX
     AAB62729;
XX
DT   03-APR-2001 (first entry)
XX
DE   Borrelia sp chimeric ospc protein SEQ ID NO: 62.
XX
KW   Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
XX
PN   WO200078966-A1.
XX
PD   28-DEC-2000.
XX
PF   19-JUN-2000; 2000WO-US16915.
XX
PR   18-JUN-1999; 99US-0140042.
XX
PI   Dattwyler RJ, Seinost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
PA   (UYNY ) UNIV NEW YORK STATE RES FOUND.
PA   (BROO-) BROOK BIOTECHNOLOGIES INC.
DR   WPI; 2001-050113/06.
DR   N-PSDB; AAF29033.
XX
PT   Compositions of OspC polypeptides from strains of Borrelia which cause
PT   Lyme disease are used to immunize animals and detect immune responses
PT   to Lyme disease -
XX
PS   Claim 43; Page 125-126; 160pp; English.
XX
CC   The present invention provides compositions comprising ospC proteins and
CC   chimeric ospC proteins from members of the Borrelia genus. These may be
CC   Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC   vaccines against Borrelia infection, which is spread by ticks and leads
CC   to Lyme disease.
XX
SQ   Sequence 397 AA;
```

Query Match   82.4%; Score 1481; DB 22; Length 397;
Best Local Similarity 83.1%; Pred. No. 2e-88;
Matches 310; Conservative 23; Mismatches 34; Indels 6; Gaps 2;

```
Qy    2 ACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKAI 61
        :||||||||||||||||||||||||||||||||||||||||||||||||||:||||||
Db   25 scnnsgkgntsansadesvkgpnlteiskkitdsnavllavkeveallssideiaakai 84

Qy   62 GKKI-GNNGLEANQSKNTLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNK 120
        |||| |||:||||||  ||  ||:|| ||||||| :||||||||||:|| |||| ||||
Db   85 gkkihqgngldteynhngsllagayaistlikqkldglkneglkekidaakkcsetftnk 144

Qy  121 LKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKE 180
```

Right column

```
Db  145 lkehtcdlgkegvtdadakeailktngtktkgaee1gklfesvev1skaakemlansvke 204

Qy  181 LTSPIVHG------NNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLL 235
        |||:|:|       ||||||||||||||||||||||||||||||||||||||||||:||||
Db  205 ltspvvaespamvnnsgkdgntsansadesvkgpnlteiskkitesnavvlavke1et11 264

Qy  236 ASIDELATKAIGKKIQONGGLAVEAGHNGTLLLAGAYTISKLIITQKLIDGLKNSEKLKEKIE 295
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  265 asidelatkaikqngglaveaghngtllagaytiskiitqkldglknseklkekie 324

Qy  296 NAKKCSEDFTKLEGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLAK 355
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  325 nakkcsedftkklegehaqlgienvtdenakkailitdaakdkgaaeleklfkavenlak 384

Qy  356 AAKEMLANSVKEL 368
        |||||||||||||
Db  385 aakemlansvkel 397
```

RESULT 10

```
ID   AAB62716 standard; Protein; 369 AA.
XX
AC   AAB62716;
XX
DT   03-APR-2001 (first entry)
XX
DE   Borrelia sp chimeric ospc protein SEQ ID NO: 36.
XX
KW   Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
XX
PN   WO200078966-A1.
XX
PD   28-DEC-2000.
XX
PF   19-JUN-2000; 2000WO-US16915.
XX
PR   18-JUN-1999; 99US-0140042.
XX
PI   Dattwyler RJ, Seinost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
PA   (UYNY ) UNIV NEW YORK STATE RES FOUND.
PA   (BROO-) BROOK BIOTECHNOLOGIES INC.
DR   WPI; 2001-050113/06.
DR   N-PSDB; AAF29020.
XX
PT   Compositions of OspC polypeptides from strains of Borrelia which cause
PT   Lyme disease are used to immunize animals and detect immune responses
PT   to Lyme disease -
XX
PS   Claim 43; Page 94-95; 160pp; English.
XX
CC   The present invention provides compositions comprising ospC proteins and
CC   chimeric ospC proteins from members of the Borrelia genus. These may be
CC   Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC   vaccines against Borrelia infection, which is spread by ticks and leads
CC   to Lyme disease.
XX
SQ   Sequence 369 AA;
```

Query Match   82.3%; Score 1480; DB 22; Length 369;
Best Local Similarity 83.8%; Pred. No. 2.2e-88;
Matches 310; Conservative 25; Mismatches 31; Indels 4; Gaps 4;

```
Qy    2 ACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKAI 61
        :||||||||||||||||||||||||||||||||||||||||||||||||||:||||||
Db    1 acnnsgkdgntsansadesvkgpnlteiskkitdsnavllavkeveallssidela-kai 59
```

```
RESULT 11
AAB62739
ID   AAB62739 standard; Protein; 400 AA.
XX
AC   AAB62739;
XX
DT   03-APR-2001  (first entry)
XX
DE   Borrelia sp chimeric ospc protein SEQ ID NO: 82.
XX
KW   Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
XX
PN   WO200078966-A1.
XX
PD   28-DEC-2000.
XX
PF   19-JUN-2000; 2000WO-US016915.
XX
PR   18-JUN-1999;   99US-0140042.
XX
PA   (UYNY ) UNIV NEW YORK STATE RES FOUND.
PA   (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI   Dattwyler RJ, Selnost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
PT   Compositions of OspC polypeptides from strains of Borrelia which cause
PT   Lyme disease are used to immunize animals and detect immune responses
PT   to Lyme disease.
XX
DR   WPI; 2001-050113/06.
DR   N-PSDB; AAF29043.
XX
CC   Claim 43; Page 153; 160pp; English.
XX
PS   Claim 43; Page 153; 160pp; English.
XX
CC   The present invention provides compositions comprising ospc proteins and
CC   chimeric ospc proteins from members of the Borrelia genus. These may be
CC   Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC   vaccines against Borrelia infection, which is spread by ticks and leads
CC   to Lyme disease.
XX
SQ   Sequence 400 AA;

  Query Match             82.3%;    Score 1479.5;  DB 22;  Length 400;

Qy        62 GKKIGNNG-LEANQSKNTSLLSGAYAISDLIAEKLNVLKNEE-LKEKIDTAKQCCSTEFTN 119
              :|:||||| |:| : |||| :::||||  |||| | | ||||| |:||| ||:
Db        60 gkkikndgsldneanrneslagaytistltcqklskIngseglkekiaaakkcseefst 119

Qy       120 KLKSEHAVLGLDNLTTDDNAQRAILLKKH-ANKDKGAAELEKLFKRAVENLSKAAQDTLKNAV 178
              ||| :||:::|||:|| : ||||::: ||||:|: |||||||||:|||| |:|
Db       119 klkdnhaqlglggvtdenakkailkanaagkdkgveeleklsgslesIskaakemlansv 178

Qy       179 KELTSPIVHGHGNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASI 238
              ::::: |:||||:|||||||||||||||||||| | ::|||||||||||||||:|:|
Db       179 kIkdnhaqlglggvtdenakkailkanaagkdkgveeleklsgslesIskaakemlansv 238

Qy       180 keltspvvhgnnsgkdgntsansadesvkgpnlteiskkitesnavvlavkeietllasi 239
              |||:||:||||:||||||||||||||||||||||| ||||||||||||||||||||||||
Db       239 DELATKAIGKKIQQNGGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKEKIENAK 298

Qy       240 delatkaigkkiqqnggIaveaghngtIlagaytiskIitqkIdgIknsekIkekienak 299
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       299 KCSEDFTKKLEGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKRAVENLAKAAK 358

Qy       299 KCSEDFTKKLEGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKRAVENLAKAAK 358
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       359 kcsedftkklegehaqlgienvtdenakkailitdaakdkgaaeleklfkavenlakaak 359

Qy       359 EMLANSVKEL 368
              ||||||||||
Db       360 emlansvkel 369
```

```
RESULT 12
AAB62731
ID   AAB62731 standard; Protein; 393 AA.
XX
AC   AAB62731;
XX
DT   03-APR-2001  (first entry)
XX
DE   Borrelia sp chimeric ospc protein SEQ ID NO: 66.
XX
KW   Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
XX
PN   WO200078966-A1.
XX
PD   28-DEC-2000.
XX
PF   19-JUN-2000; 2000WO-US016915.
XX
PR   18-JUN-1999;   99US-0140042.
XX
PA   (UYNY ) UNIV NEW YORK STATE RES FOUND.
PA   (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI   Dattwyler RJ, Selnost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
PT   Compositions of OspC polypeptides from strains of Borrelia which cause
PT   Lyme disease are used to immunize animals and detect immune responses
PT   to Lyme disease.
XX
DR   WPI; 2001-050113/06.
DR   N-PSDB; AAF29035.
XX
PS   Claim 43; Page 131; 160pp; English.
XX
CC   The present invention provides compositions comprising ospc proteins and
CC   chimeric ospc proteins from members of the Borrelia genus. These may be
CC   Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC   vaccines against Borrelia infection, which is spread by ticks and leads

  Best Local Similarity   82.4%;    Conservative 23;  Matches 310;  Mismatches 34;  Indels  9;  Gaps  2;
  Matches 310;  Conservative 23;  Pred. No. 2.6e-88;

Qy         2 ACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKAI 61
              | :|  |||: ||||||||||||||||||||||||||||||||:|| |:|:|||||||
Db        25 scnnsgkdgntsansadesvkgpnIteiskkitdsnavIIavkeveaIIssideIaakai 84

Qy        62 GKKI-GNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCCSTEFTNK 120
              |||| ||: |  : |||| :::|||| ||:|||| | | |||| |:|||  ||:
Db        85 gkkihgqngldteynhgsllagayaistlikgkIdglkneglkekidaakkcsetftnk 144

Qy       121 LKSEHAVLGLDNLTDDNAQRILKKHANKDKGAAELEKLFKRAVENLSKAAQDTLKNAVKE 180
              ||| :||:::|||:|| : ||:::: ||:|| ||||||||||:||||  |:|
Db       145 IkehtdlgkegvtdadakeaiIktngtkkgaeelfesvevIskaakemIansvke 204

Qy       181 LTSPIVHG-------NNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIE 232
              |||:|:||       ||:||||||||||||||||||||||||||||||||||||||||
Db       205 ItspvvaespkkpsmvnnsgkdgntsansadesvkgpnIteiskkitesnavvIavkeie 264

Qy       233 TLLASIDELATKAIGKKIQQNGGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKE 292
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       265 tIIasideIatkaigkkiqqnggIaveaghngtIIagaytiskIitqkIdgIknsekIke 324

Qy       293 KIENAKKCSEDFTKKLEGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVEN 352
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       325 kienakkcsedftkkIegehaqIgienvtdenakkaiIitdaakdkgaaeIekIfkaven 384

Qy       353 LAKAAKEMLANSVKEL 368
              ||||||||||||||||
Db       385 lakaakemlansvkel 400
```

RESULT 13

ID   AAB62712 standard; protein; 378 AA.
XX
AC   AAB62712;
XX
DT   03-APR-2001 (first entry)
XX
DE   Borrelia sp chimeric ospC protein SEQ ID NO: 28.
XX
KW   Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
XX
PN   WO200078966-A1.
XX
PD   28-DEC-2000.
XX
PR   18-JUN-1999; 99US-0140042.
XX
PF   19-JUN-2000; 2000WO-US16915.
XX
PI   Dattwyler RJ, Selnost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
PA   (UYNY) UNIV NEW YORK STATE RES FOUND.
PA   (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
DR   WPI; 2001-050113/06.
DR   N-PSDB; AAF29016.
XX
PT   Compositions of OspC polypeptides from strains of Borrelia which cause
PT   Lyme disease are used to immunize animals and detect immune responses
PT   to Lyme disease -
XX
CC   to Lyme disease.
XX
SQ   Sequence   393 AA;

Query Match   82.1%; Score 1477; DB 22; Length 393;
Best Local Similarity   83.5%; Pred. No. 3.7e-88;
Matches 309; Conservative 26; Mismatches 31; Indels   4; Gaps   4;

Qy   2 ACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKAI 61
         :||||||||| ||||||||||||||||||| ||||||||||||||:|:|||||||| |||
Db   25 scnnsgkdgntsansadesvkgpnlteiskkitdsnavllavkeveeallssidela-kai 83

Qy   62 GKKIGNNG-LEANQSKNTSLLSGAYAISDLIAEKLNVLKNEE-LKEKIDTAKQCSTEFTN 119
         ||: ::|| :||||||||||||||||||:||||||||:||| ||:|||||||||| ||
Db   84 gkkikndgsldneanrneslla gaytistlitqklsklngseglkekiaaakkcseefst 143

Qy   120 KLKSEHAVIGLDNLTDDNAQRAILKH-ANKDKGAAELEKLFKAVENLSKAAQDTLKNAV 178
         |||||||||||||||||||||||||| ||:|||||||||||||||:||:||:|||||||:
Db   144 klkdnhaqlgiggvtdenakkailkanaagkdkgveeleklsgsleslskaakemlansv 203

Qy   179 KELTSPIVHGNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASI 238
         |:|||||||||||||||||||||||||||||||||||||||||||||||||:||||||||
Db   204 keltspvvhgnnsgkdgntsansadesvkgpnlteiskkitesnavvlavkeletllasi 263

Qy   239 DELATKAIGKKIQQNGGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKEIENAK 298
         ||||||||||||||||| |||||||||||||||||||| ||||||||||||||||||||
Db   264 delatkaigkkiqqnggtlaveaghngtllagaytiskliitqkidglknseklkeienak 323

Qy   299 KCSEDFTKKLEGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLAKAAK 358
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   324 kcsedftkklegehaqlgienvtdenakkailitdaakdkgaaeleklfkavenlakaak 383

Qy   359 EMLANSVKEL 368
         ||||||||||
Db   384 emlansvkel 393

RESULT 14

ID   AAB62710 standard; protein; 374 AA.
XX
AC   AAB62710;
XX
DT   03-APR-2001 (first entry)
XX
DE   Borrelia sp chimeric ospc protein SEQ ID NO: 24.
XX
KW   Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
XX
PN   WO200078966-A1.
XX
PD   28-DEC-2000.
XX
PR   18-JUN-1999; 99US-0140042.
XX
PF   19-JUN-2000; 2000WO-US16915.
XX
PI   Dattwyler RJ, Selnost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
PA   (UYNY) UNIV NEW YORK STATE RES FOUND.
PA   (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
DR   WPI; 2001-050113/06.
XX
PS   Claim 43; Page 83-84; 160pp; English.
XX
CC   The present invention provides compositions comprising ospc proteins and
CC   chimeric ospc proteins from members of the Borrelia genus. These may be
CC   Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC   vaccines against Borrelia infection, which is spread by ticks and leads
CC   to Lyme disease.
XX
SQ   Sequence   378 AA;

Query Match   69.5%; Score 1250; DB 22; Length 378;
Best Local Similarity   70.3%; Pred. No. 1.7e-73;
Matches 265; Conservative 43; Mismatches 59; Indels 10; Gaps   3;

Qy   1 MACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKA 60
         || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 macnnsgkdgntsansadesvkgpnlteiskkitdsnavllavkeveeallssidelaaka 60

Qy   61 IGKKI-GNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTN 119
         ||||| |: :||||||||||||||||||||:|||||||||: |||||||||||| ||
Db   61 igkkihgnngldtennhngslla gayaistliikqkidglkneglkekidaakkcsetftn 120

Qy   120 KLKSEHAVIGLDNLTDDNAQRAILKKIHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVK 179
         |||||||||||||||||||||||| : || : ||||||||||||||||||: | ||||||
Db   121 klkehtdlgkegvtdadakeailkangtktkgaaelgkifesvevlskaakemlansvk 179

Qy   180 ELTSPIVHG-----------NNSKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEI 231
         ||||:|:| :||||||||||||||||||||||||||||||||||||||:||
Db   181 eltspvvaespkkpsmvnnsgkdgntsansadesvkgpnlteiskkitesnavvlavkev 240

Qy   232 ETLLASIDELATKAIGKKIQQNGGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLK 291
         |||||||||||||||| ||||:|||||||||||||||||||||||||||||||||||| |
Db   241 etlltsidela-kaigkkindvsidneadhngslisgaylisnlitkkisaidsgelk 299

Qy   292 EKIENAKKCSEDFTKKLEGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVE 351
         :|| : || ||||||||||| |||:| ||:||||||||||:|:| ||||||||||| |
Db   300 aeiekakkcseeftaklkgehtdlgkegvtddnakkailktnndktkgadeleklfesvk 359

Qy   352 NLAKAKEMLANSVKEL 368
         |: || |||||||||
Db   360 nlskaakemlnsvkel 376

ID   AAB62728 standard; Protein; 398 AA.
XX
AC   AAB62728;
XX
DT   03-APR-2001 (first entry)
XX
DE   Borrelia sp chimeric ospC protein SEQ ID NO: 60.
XX
KW   Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
XX
PN   WO200078966-A1.
XX
PD   28-DEC-2000.
XX
PF   19-JUN-2000; 2000WO-US16915.
XX
PR   18-JUN-1999; 99US-0140042.
XX

DR   N-PSDB; AAF29014.
XX
PT   Lyme disease are used to immunize animals and detect immune responses
PT   Compositions of OspC polypeptides from strains of Borrelia which cause
PT   to Lyme disease -
XX
PS   Claim 43; Page 78-79; 160pp; English.
XX
CC   The present invention provides compositions comprising ospC proteins and
CC   chimeric ospC proteins from members of the Borrelia genus. These may be
CC   Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC   vaccines against Borrelia infection, which is spread by ticks and leads
CC   to Lyme disease.
XX
SQ   Sequence   374 AA;

Query Match   69.4%; Score 1248.5; DB 22; Length 374;
Best Local Similarity 71.0%; Pred. No. 2.1e-73;
Matches 265; Conservative 42; Mismatches 59; Indels 7; Gaps 3;

Qy   2 ACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKAI   61
             ||| ||||||:|||||||||||||||||||||||:||||:||||||||||:||:|:||:||
Db   1 acnnsgkdgntsansadesvkgpnlteiskkitdsnavllavkeveallssideiaakai   60

Qy   62 GKKI-GNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNK 120
             |||| ||||: ||:::||||||||||:|||:|:|||||||:||||||||:||||:|||||
Db   61 gkkihqnngidteynhngsllagayaistilkqkldglkneglkekidaakkcsetftnk 120

Qy   121 LKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKE 180
             || :| ||:|: :|.:||: |||:||||||||||||||||||||||||||||||:||||
Db   121 lkekhtdlgkegvtdadakeailktngtktkgaeelgklfesvevlskaakemlansvke 180

Qy   181 LTSPIYHG-----NNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLL 235
             ||||: | |||||||||||||||||||||||||||||||||||||||||:||||||
Db   181 ltspvvaespamvnnsgkdgntsansadesvkgpnlteiskkitesnavvlavkevetll 240

Qy   236 ASIDELATKAIGKKIQONGGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKEKIE 295
             :||||| |||||||:|:| |||:|| |||||||||||||:|||:|||:|||||:||||||
Db   241 tsidela-kaigkkikndvsldneadhngslisgaylisnltkkisaikdsgelkaele 299

Qy   296 NAKKCSEDFTKLEGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLAK 355
             :||||||:||:|:|||:.||||:|:.||||| |.|||:||:|.|||||||:|||:|||:|
Db   300 kakkcseeftaklkgehtdlgkegvtddnakkalktnndktkgadelekifesvknlsk 359

Qy   356 AAKEMLANSVKEL 368
             ||||||:||||||
Db   360 aakemltnsvkel 372

PA   (UNY) UNIV NEW YORK STATE RES FOUND.
XX
PA   (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI   Dattwyler RJ, Seinost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
DR   WPI; 2001-050113/06.
DR   N-PSDB; AAF29032.
XX
PT   Lyme disease are used to immunize animals and detect immune responses
PT   Compositions of OspC polypeptides from strains of Borrelia which cause
PT   to Lyme disease -
XX
PS   Claim 43; Page 123; 160pp; English.
XX
CC   The present invention provides compositions comprising ospC proteins and
CC   chimeric ospC proteins from members of the Borrelia genus. These may be
CC   Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC   vaccines against Borrelia infection, which is spread by ticks and leads
CC   to Lyme disease.
XX
SQ   Sequence   398 AA;

Query Match   69.3%; Score 1245.5; DB 22; Length 398;
Best Local Similarity 70.8%; Pred. No. 3.6e-73;
Matches 264; Conservative 43; Mismatches 59; Indels 7; Gaps 3;

Qy   2 ACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKAI   61
             :|||||||||:|||||||||||||||||||||||:||||:||||||||||:||:|:||:||
Db   25 scnnsgkdgntsansadesvkgpnlteiskkitdsnavllavkeveallssideiaakai   84

Qy   62 GKKI-GNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNK 120
             |||| ||||: ||:::||||||||||:|||:|:|||||||:||||||||:||||:|||||
Db   85 gkkihqnngidteynhngsllagayaistilkqkldglkneglkekidaakkcsetftnk 144

Qy   121 LKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKE 180
             || :| ||:|: :|.:||: |||:||||||||||||||||||||||||||||||:||||
Db   145 lkekhtdlgkegvtdadakeailktngtktkgaeelgklfesvevlskaakemlansvke 204

Qy   181 LTSPIYHG-----NNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLL 235
             ||||: | |||||||||||||||||||||||||||||||||||||||||:||||||
Db   205 ltspvvaespamvnnsgkdgntsansadesvkgpnlteiskkitesnavvlavkevetll 264

Qy   236 ASIDELATKAIGKKIQONGGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKEKIE 295
             :||||| |||||||:|:| |||:|| |||||||||||||:|||:|||:|||||:||||||
Db   265 tsidela-kaigkkikndvsldneadhngslisgaylisnltkkisaikdsgelkaele 323

Qy   296 NAKKCSEDFTKLEGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLAK 355
             :||||||:||:|:|||:.||||:|:.||||| |.|||:||:|.|||||||:|||:|||:|
Db   324 kakkcseeftaklkgehtdlgkegvtddnakkalktnndktkgadelekifesvknlsk 383

Qy   356 AAKEMLANSVKEL 368
             ||||||:||||||
Db   384 aakemltnsvkel 396

Search completed: March 18, 2002, 09:54:34
Job time: 332 sec

OM protein - protein search, using sw model

```
                    GenCore version 4.5
                    Copyright (c) 1993 - 2000 Compugen Ltd.

Run on:             March 18, 2002, 10:08:45 ; Search time 621.2 Seconds
                                                (without alignments)
                                                164.485 Million cell updates/sec

Title:              US-09-596-746A-42
Perfect score:      1798
Sequence:           1 MACNNSGKDGNASANSADES..........AVENLAKAAKEMLANSVKEL 368

Scoring table:
Searched:           BLOSUM62
                    Gapop 10.0 , Gapext 0.5
Total number of hits satisfying chosen parameters:     3148936
                    3148936 seqs, 27657034 residues

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing:    Minimum Match 0%
                    Maximum Match 100%
                    Listing first 45 summaries

Database :          Pending_Patents_AA_Main:*
                    1: /cgn2_6/ptodata/2/paa/PCTUS_COMB.pep:*
                    2: /cgn2_6/ptodata/2/paa/US06_COMB.pep:*
                    3: /cgn2_6/ptodata/2/paa/US07_COMB.pep:*
                    4: /cgn2_6/ptodata/2/paa/US080_COMB.pep:*
                    5: /cgn2_6/ptodata/2/paa/US081_COMB.pep:*
                    6: /cgn2_6/ptodata/2/paa/US082_COMB.pep:*
                    7: /cgn2_6/ptodata/2/paa/US083_COMB.pep:*
                    8: /cgn2_6/ptodata/2/paa/US084_COMB.pep:*
                    9: /cgn2_6/ptodata/2/paa/US085_COMB.pep:*
                    10: /cgn2_6/ptodata/2/paa/US086_COMB.pep:*
                    11: /cgn2_6/ptodata/2/paa/US087_COMB.pep:*
                    12: /cgn2_6/ptodata/2/paa/US088_COMB.pep:*
                    13: /cgn2_6/ptodata/2/paa/US089_COMB.pep:*
                    14: /cgn2_6/ptodata/2/paa/US090_COMB.pep:*
                    15: /cgn2_6/ptodata/2/paa/US091_COMB.pep:*
                    16: /cgn2_6/ptodata/2/paa/US092_COMB.pep:*
                    17: /cgn2_6/ptodata/2/paa/US093_COMB.pep:*
                    18: /cgn2_6/ptodata/2/paa/US094_COMB.pep:*
                    19: /cgn2_6/ptodata/2/paa/US095_COMB.pep:*
                    20: /cgn2_6/ptodata/2/paa/US096_COMB.pep:*
                    21: /cgn2_6/ptodata/2/paa/US097_COMB.pep:*
                    22: /cgn2_6/ptodata/2/paa/US098_COMB.pep:*
                    23: /cgn2_6/ptodata/2/paa/US099_COMB.pep:*
                    24: /cgn2_6/ptodata/2/paa/US60_COMB.pep:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Match Length	DB ID	Description
1	1798	100.0	368	19	US-09-596-746A-42	Sequence 42, Appl
2	1793	99.7	367	19	US-09-596-746A-76	Sequence 76, Appl
3	1790	99.6	391	19	US-09-596-746A-74	Sequence 74, Appl
4	1790	99.6	392	19	US-09-596-746A-74	Sequence 74, Appl
5	1562.5	86.9	369	19	US-09-596-746A-40	Sequence 40, Appl
6	1557.5	86.6	368	19	US-09-596-746A-40	Sequence 40, Appl
7	1554.5	86.5	392	19	US-09-596-746-72	Sequence 72, Appl
8	1554.5	86.5	393	19	US-09-596-746A-72	Sequence 72, Appl
9	1538	85.5	368	19	US-09-596-746-38	Sequence 38, Appl

10	1538	85.5	368	19	US-09-596-746A-38	Sequence 38, Appl
11	1530	85.1	391	19	US-09-596-746A-76	Sequence 76, Appl
12	1530	85.1	392	19	US-09-596-746A-76	Sequence 76, Appl
13	1489	82.8	374	19	US-09-596-746A-26	Sequence 26, Appl
14	1485.5	82.6	374	19	US-09-596-746A-30	Sequence 30, Appl
15	1485.5	82.6	377	19	US-09-596-746A-30	Sequence 30, Appl
16	1485	82.6	370	19	US-09-596-746A-36	Sequence 36, Appl
17	1484	82.5	373	19	US-09-596-746A-26	Sequence 26, Appl
18	1481	82.4	397	19	US-09-596-746A-62	Sequence 62, Appl
19	1481	82.4	398	19	US-09-596-746A-62	Sequence 62, Appl
20	1480	82.3	369	19	US-09-596-746A-36	Sequence 36, Appl
21	1479.5	82.3	400	19	US-09-596-746A-82	Sequence 82, Appl
22	1479.5	82.3	401	19	US-09-596-746A-82	Sequence 82, Appl
23	1477	82.1	393	19	US-09-596-746A-66	Sequence 66, Appl
24	1477	82.1	394	19	US-09-596-746A-66	Sequence 66, Appl
25	1253.5	69.7	375	19	US-09-596-746A-24	Sequence 24, Appl
26	1250	69.5	378	19	US-09-596-746A-28	Sequence 28, Appl
27	1250	69.5	371	19	US-09-596-746A-34	Sequence 34, Appl
28	1249.5	69.5	378	19	US-09-596-746A-28	Sequence 28, Appl
29	1248.5	69.4	371	19	US-09-596-746A-34	Sequence 34, Appl
30	1245.5	69.3	398	19	US-09-596-746A-60	Sequence 60, Appl
31	1245.5	69.3	399	19	US-09-596-746A-60	Sequence 60, Appl
32	1244.5	69.2	370	19	US-09-596-746A-34	Sequence 34, Appl
33	1245.5	69.2	401	19	US-09-596-746A-80	Sequence 80, Appl
34	1244	69.2	402	19	US-09-596-746A-80	Sequence 80, Appl
35	1244	69.1	401	19	US-09-596-746A-64	Sequence 64, Appl
36	1241.5	69.0	395	19	US-09-596-746A-64	Sequence 64, Appl
37	1241.5	69.0	370	19	US-09-596-746A-32	Sequence 32, Appl
38	1225	68.1	370	19	US-09-596-746A-32	Sequence 32, Appl
39	1220	67.9	393	19	US-09-596-746A-68	Sequence 68, Appl
40	1217	67.7	394	19	US-09-596-746A-68	Sequence 68, Appl
41	1217	67.7	393	19	US-09-596-746A-68	Sequence 68, Appl
42	1167.5	64.9	560	19	US-09-596-746A-52	Sequence 52, Appl
43	1167.5	64.9	560	19	US-09-596-746A-52	Sequence 52, Appl
44	1164.5	64.8	560	19	US-09-596-746A-56	Sequence 56, Appl
45	1159.5	64.5	385	19	US-09-596-746A-56	Sequence 56, Appl
46	1156.5	64.3	408	19	US-09-596-746-78	Sequence 78, Appl

ALIGNMENTS

```
RESULT     1
US-09-596-746A-42

Sequence 42, Application US/0959674A
GENERAL INFORMATION:
APPLICANT:          Dattwyler, Raymond J.
APPLICANT:          Seinost, Gerald
APPLICANT:          Dykhuizen, Daniel
APPLICANT:          Luft, Benjamin J.
APPLICANT:          Maria J.C. Gomes-Solecki
TITLE OF INVENTION: Groups of Borrelia burgdorferi and
TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
FILE REFERENCE:     2631.1002-001
CURRENT APPLICATION NUMBER: US/09/596,746A
CURRENT FILING DATE: 2000-06-19
PRIOR APPLICATION NUMBER: US 60/140,042
PRIOR FILING DATE: 1999-06-18
NUMBER OF SEQ ID NOS: 84
SOFTWARE:           FastSEQ for Windows Version 4.0
SEQ ID NO 42
    LENGTH: 368
    TYPE: PRT
    ORGANISM: Artificial Sequence
    FEATURE:
    OTHER INFORMATION: OspC Chimera
US-09-596-746A-42
```

```
Qy

Query Match             100.0%;    Score 1798;  DB 19;    Length 368;
Best Local Similarity   100.0%;    Pred. No. 5.6e-124;
    Matches  368;   Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1 MACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKA 60
```

Db   1 MACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKA   60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   61 IGKKIGNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNK 120
Db   61 IGKKIGNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNK 120

Qy 121 LKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKE 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 LKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKE 180

Qy 181 LTSPIVHGNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDE 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 LTSPIVHGNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDE 240

Qy 241 LATKAIGKKIQONGGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKEKIENAKKC 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 241 LATKAIGKKIQONGGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKEKIENAKKC 300

Qy 301 SEDFTKKLEGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLAKAAKEM 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 301 SEDFTKKLEGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLAKAAKEM 360

Qy 361 LANSVKEL 368
         ||||||||
Db 361 LANSVKEL 368

RESULT 2
US-09-596-746-42
Sequence 42, Application US/0959674 6
GENERAL INFORMATION:
  APPLICANT: Dattwyler, Raymond J.
  APPLICANT: Seinost, Gerald
  APPLICANT: Dykhuizen, Daniel
  APPLICANT: Luft, Benjamin J.
  APPLICANT: Maria J.C. Gomes-Solecki
  TITLE OF INVENTION: Groups of Borrelia burgdorferi and
  TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
  FILE REFERENCE: 2631.1002-001
  CURRENT APPLICATION NUMBER: US-09/596,746
  CURRENT FILING DATE: 2000-06-16
  PRIOR APPLICATION NUMBER: US 60/140,042
  PRIOR FILING DATE: 1999-06-18
  NUMBER OF SEQ ID NOS: 84
  SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO: 42
  LENGTH: 367
  TYPE: PRT
  ORGANISM: Artificial Sequence
  FEATURE:
  OTHER INFORMATION: OspC Chimera
US-09-596-746-42

Query Match   99.7%; Score 1793; DB 19; Length 367;
Best Local Similarity 100.0%; Pred. No. 1.3e-123;
Matches 367; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 181 TSPIVHGNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDEL 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 361 ANSVKEL 367

Db 181 TSPIVHGNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDEL 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 242 ATKAIGKKIQONGGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKEKIENAKKCS 301
Db 241 ATKAIGKKIQONGGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKEKIENAKKCS 300

Qy 302 EDFTKKLEGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLAKAAKEML 361
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 301 EDFTKKLEGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLAKAAKEML 360

Qy 362 ANSVKEL 368
         |||||||
Db 361 ANSVKEL 367

RESULT 3
US-09-596-746-74
Sequence 74, Application US/0959674 6
GENERAL INFORMATION:
  APPLICANT: Dattwyler, Raymond J.
  APPLICANT: Seinost, Gerald
  APPLICANT: Dykhuizen, Daniel
  APPLICANT: Luft, Benjamin J.
  APPLICANT: Maria J.C. Gomes-Solecki
  TITLE OF INVENTION: Groups of Borrelia burgdorferi and
  TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
  FILE REFERENCE: 2631.1002-001
  CURRENT APPLICATION NUMBER: US-09/596,746
  CURRENT FILING DATE: 2000-06-16
  PRIOR APPLICATION NUMBER: US 60/140,042
  PRIOR FILING DATE: 1999-06-18
  NUMBER OF SEQ ID NOS: 84
  SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO: 74
  LENGTH: 391
  TYPE: PRT
  ORGANISM: OspC Chimera
US-09-596-746-74

Query Match   99.6%; Score 1790; DB 19; Length 391;
Best Local Similarity 99.7%; Pred. No. 2.4e-123;
Matches 366; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy   2 ACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKAI   61
         :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   25 SCNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKAI   84

Qy   62 GKKIGNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNKL 121
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   85 GKKIGNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNKL 144

Qy 122 KSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKEL 181
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 145 KSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKEL 204

Qy 182 TSPIVHGNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDEL 241
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 205 TSPIVHGNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDEL 264

Qy 242 ATKAIGKKIQONGGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKEKIENAKKCS 301
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 265 ATKAIGKKIQONGGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKEKIENAKKCS 324

Qy 302 EDFTKKLEGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLAKAAKEML 361
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 325 EDFTKKLEGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLAKAAKEML 384

Qy 362 ANSVKEL 368
         |||||||
Db 385 ANSVKEL 391

RESULT 4
US-09-596-746A-74
Sequence 74, Application US/0959674A
GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seinost, Gerald
; APPLICANT: Dykhuizen, Danial
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746A
; CURRENT FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 74
; LENGTH: 392
; TYPE: PRT
; ORGANISM: ospc Chimera
US-09-596-746A-74

Query Match   99.6%;   Score 1790;   DB 19;   Length 392;
Best Local Similarity   99.7%;   Pred. No. 2.4e-123;
Matches 366;   Conservative   1; Mismatches   0; Indels   0; Gaps   0;

Qy   2 ACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKAI 61
          :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   26 SCNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKAI 85

Qy   62 GKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNKL 121
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   86 GKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNKL 145

Qy   122 KSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKEL 181
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   146 KSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKEL 205

Qy   182 TSPIVHGNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDEL 241
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   206 TSPIVHGNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDEL 265

Qy   242 ATKAIGKKIQONGGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKEKIENAKKCS 301
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   266 ATKAIGKKIQONGGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKEKIENAKKCS 325

Qy   302 EDFTKKLEGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLAKAAKEML 361
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   326 EDFTKKLEGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLAKAAKEML 385

Qy   362 ANSVKEL 368
          |||||||
Db   386 ANSVKEL 392

RESULT 5
US-09-596-746A-40
Sequence 40, Application US/0959674A
GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seinost, Gerald
; APPLICANT: Dykhuizen, Danial
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: 2000-06-19
; CURRENT FILING DATE: US/09/596,746A
; PRIOR APPLICATION NUMBER: US 60/140,042

Query Match   86.6%;   Score 1557.5;   DB 19;   Length 368;

; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 40
; LENGTH: 369
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: OspC Chimera
US-09-596-746A-40

Query Match   86.9%;   Score 1562.5;   DB 19;   Length 369;
Best Local Similarity   87.5%;   Pred. No. 1.3e-105;
Matches 322;   Conservative   20; Mismatches   25; Indels   1; Gaps   1;

Qy   1 MACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKA 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKA 60

Qy   61 IGKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNK 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 IGKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNK 120

Qy   121 LKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKE 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 LKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKE 180

Qy   181 LTSPIVHGNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDE 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||
Db   181 LTSPIVHGNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDE 240

Qy   241 LATKAIGKKIQONGGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKEKIENAKKC 300
          |||||||||||| | |||:||||||||||||||||||||||||:||||||||||||||||
Db   241 LA-KAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAEIEKAKKC 299

Qy   301 SEDFTKKLEGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLAKAAKEM 360
          |||||||||||:||||||||||||::|||:|||||||||||||||||||||||||||||
Db   300 SEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFFESVKNLSRAAKEM 359

Qy   361 LANSVKEL 368
          | |||||||
Db   360 LTNSVKEL 367

RESULT 6
US-09-596-746-40
Sequence 40, Application US/0959674A
GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seinost, Gerald
; APPLICANT: Dykhuizen, Danial
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746
; CURRENT FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 40
; LENGTH: 368
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: OspC Chimera.
US-09-596-746A-40

Query Match   86.6%;   Score 1557.5;   DB 19;   Length 368;

```
Best Local Similarity   87.5%;
Matches 321; Conservative 20; Mismatches 25; Indels 1; Gaps 1;

Qy    2 ACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKAI 61
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 ACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKAI 60

Qy   62 GKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNKL 121
        |:||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 GKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNKL 120

Qy  122 KSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKEL 181
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 KSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKEL 180

Qy  182 TSPIVHGNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDEL 241
        |||||||||||||||||:||||||||||||||||||||||||||||||:|||||||:||||
Db  181 TSPIVHGNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLTSIDEL 240

Qy  242 ATKAIGKKIQONGGLAVEAGHNGTLLAGAYTISKLIITQKLDGLKNSEKLKEKIENAKKCS 301
        |:|||||||:|:||| || ||||:||||||||:|||:|:|||||||||||| | ||||
Db  241 A-KAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAEIEKAKKCS 299

Qy  302 EDFTKKLEGEHAQLGIENVTDENAKKAILITDAAAKDKGAAELEKLFKAVENLAKAAKEML 361
        |:|| ||:||| |||||||||||:||:|::||| |:|||||||||| |:|||:||||||
Db  300 EEFTAKLKGEHTDLGKEGVTDNAKKAILKTNNDKTKGADELEKLFESVKNLSKAAKEML 359

Qy  362 ANSVKEL 368
        :|||||
Db  360 TNSVKEL 366
```

RESULT 7
US-09-596-746-72

US-09-596-746-72

```
Query Match     86.5%;   Score 1554.5;  DB 19;  Length 392;
Best Local Similarity   87.2%; Pred. No. 5.4e-106;
Matches 320; Conservative 21; Mismatches 25; Indels 1; Gaps 1;

Qy    2 ACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKAI 61
        :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   25 SCNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKAI 84

Qy   62 GKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNKL 121
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   85 GKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNKL 144

Qy  122 KSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKEL 181
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  145 KSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKEL 204

Qy  182 TSPIVHGNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDEL 241
        |||||||||||||||||:||||||||||||||||||||||||||||||:|||||||:||||
Db  205 TSPIVHGNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLTSIDEL 264

Qy  242 ATKAIGKKIQONGGLAVEAGHNGTLLAGAYTISKLIITQKLDGLKNSEKLKEKIENAKKCS 301
        |:|||||||:|:||| || ||||:||||||||:|||:|:|||||||||||| | ||||
Db  265 A-KAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAEIEKAKKCS 323

Qy  302 EDFTKKLEGEHAQLGIENVTDENAKKAILITDAAAKDKGAAELEKLFKAVENLAKAAKEML 361
        |:|| ||:||| |||||||||||:||:|::||| |:|||||||||| |:|||:||||||
Db  324 EEFTAKLKGEHTDLGKEGVTDNAKKAILKTNNDKTKGADELEKLFESVKNLSKAAKEML 383

Qy  362 ANSVKEL 368
        :|||||
Db  384 TNSVKEL 390
```

RESULT 8
US-09-596-746A-72

US-09-596-746A-72

```
Query Match     86.5%;   Score 1554.5;  DB 19;  Length 393;
Best Local Similarity   87.2%; Pred. No. 5.5e-106;
Matches 320; Conservative 21; Mismatches 25; Indels 1; Gaps 1;

Qy    2 ACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKAI 61
        :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   26 SCNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKAI 85

Qy   62 GKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNKL 121
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   86 GKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNKL 145

Qy  122 KSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKEL 181
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  146 KSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKEL 205

Qy  182 TSPIVHGNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDEL 241
        |||||||||||||||||:||||||||||||||||||||||||||||||:|||||||:||||
Db  206 TSPIVHGNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLTSIDEL 265

Qy  242 ATKAIGKKIQONGGLAVEAGHNGTLLAGAYTISKLIITQKLDGLKNSEKLKEKIENAKKCS 301
        |:|||||||:|:||| || ||||:||||||||:|||:|:|||||||||||| | ||||
Db  266 A-KAIGKKIKNDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAEIEKAKKCS 324

Qy  302 EDFTKKLEGEHAQLGIENVTDENAKKAILITDAAAKDKGAAELEKLFKAVENLAKAAKEML 361
        |:|| ||:||| |||||||||||:||:|::||| |:|||||||||| |:|||:||||||
Db  325 EEFTAKLKGEHTDLGKEGVTDNAKKAILKTNNDKTKGADELEKLFESVKNLSKAAKEML 384

Qy  362 ANSVKEL 368
        :|||||
Db  385 TNSVKEL 391
```

RESULT 9
US-09-596-746-38
Sequence 38, Application US/0959746
GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seinost, Gerald
; APPLICANT: Dykhuizen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: 2000-06-16
; CURRENT FILING DATE:
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSEQ for Windows version 4.0
; SEQ ID NO 38
; LENGTH: 368
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: OspC Chimera
US-09-596-746:38

Query Match   85.5%;   Score 1538;   DB 19;   Length 368;
Best Local Similarity 86.7%; Pred. No. 8.1e-105;
Matches 319; Conservative 20; Mismatches 27;   Indels   2;   Gaps   2;

Qy   1 MACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKA 60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKA 60

Qy   61 IGKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNK 120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 IGKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNK 120

Qy   121 LKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKE 180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 LKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKE 180

Qy   181 LTSPIVHGNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDE 240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 LTSPIVHGNNSRKDGNASTNSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDE 240

Qy   241 LATKAIGKKI-GNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKN-EELKEKIDTAKQC 298
         ||||| ||| :|:|||:||| ||||||||||:||||||||||||| ||:|||||||||
Db   241 LATKAIGKKIQGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQC 298

Qy   301 SEDFTKLEGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLAKAAKEM 360
         |:|| ||||||||||||:|||||||| ||:|:||||||||||||||||||||||:|||
Db   299 STEFTNKLKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDT 358

Qy   361 LANSVKEL 368
         | :|:||
Db   359 LKNAVKEL 366

; ●

RESULT 10
US-09-596-746A-38
Sequence 38, Application US/0959746A
GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seinost, Gerald
; APPLICANT: Dykhuizen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans

---

; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746A
; CURRENT FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 38
; LENGTH: 368
; TYPE: PRT
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: OspC Chimera
US-09-596-746A-38

Query Match   85.5%;   Score 1538;   DB 19;   Length 368;
Best Local Similarity 86.7%; Pred. No. 8.1e-105;
Matches 319; Conservative 20; Mismatches 27;   Indels   2;   Gaps   2;

Qy   1 MACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKA 60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKA 60

Qy   61 IGKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNK 120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 IGKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNK 120

Qy   121 LKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKE 180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 LKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKE 180

Qy   181 LTSPIVHGNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDE 240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 LTSPIVHGNNSRKDGNASTNSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDE 240

Qy   241 LATKAIGKKI-GNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKN-EELKEKIDTAKQC 298
         ||||| ||| :|:|||:||| ||||||||||:||||||||||||| ||:|||||||||
Db   241 LATKAIGKKIQGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQC 298

Qy   301 SEDFTKLEGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLAKAAKEM 360
         |:|| ||||||||||||:|||||||| ||:|:||||||||||||||||||||||:|||
Db   299 STEFTNKLKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDT 358

Qy   361 LANSVKEL 368
         | :|:||
Db   359 LKNAVKEL 366

; ●

RESULT 11
US-09-596-746-76
Sequence 76, Application US/0959746
GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seinost, Gerald
; APPLICANT: Dykhuizen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: 2000-06-16
; CURRENT FILING DATE:
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 76
; LENGTH: 391
; TYPE: PRT
; ORGANISM: OspC Chimera
US-09-596-746:76

```
RESULT 12
US-09-596-746A-76
Sequence 76, Application US/0959746A
GENERAL INFORMATION:
;   APPLICANT: Dattwyler, Raymond J.
;   APPLICANT: Seinost, Gerald
;   APPLICANT: Dykhuizen, Daniel
;   APPLICANT: Luft, Benjamin J.
;   APPLICANT: Maria J.C. Gomes-Solecki
;   TITLE OF INVENTION: Groups of Borrelia burgdorferi and
;   TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
;   FILE REFERENCE: 2631.1002-001
;   CURRENT APPLICATION NUMBER: US-09/596,746A
;   CURRENT FILING DATE: 2000-06-19
;   PRIOR APPLICATION NUMBER: US 60/140,042
;   PRIOR FILING DATE: 1999-06-18
;   NUMBER OF SEQ ID NOS: 84
;   SOFTWARE: FastSEQ for Windows Version 4.0
;   SEQ ID NO 76
;   LENGTH: 392
;   TYPE: PRT
;   ORGANISM: ospc Chimera
US-09-596-746A-76

Query Match             85.1%;  Score 1530;  DB 19;  Length 392;
Best Local Similarity   86.4%;  Pred. No. 3.5e-104;
Matches   317;  Conservative   21;  Mismatches   27;  Indels    2;  Gaps    2;

Qy          2  ACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKAI  61
                :||||||||||||||||||||||||||||||||||||||||||||||||||:||||||||
Db         26  SCNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKAI  85

Qy         62  GKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNKL 121
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         86  GKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNKL 145

Qy        122  KSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKEL 181
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145  KSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKEL 204

Qy        182  TSPIVHGNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDEL 241
               ||||||||||| :|| |||||||||||||||||||||||||||||||||||||||||||||
Db        205  TSPIVHGNNSRDGNASTNSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDEL  264

Qy        242  ATKAIGKKIQQNGGLAVEAGHNGTLLAGAYTISKLLTQKLDGLKNSEKLKEKIENAKKCS 301
               ||||||||| : ||| : | : |||  |:|| ||||  |||  :|||||:||| ||
Db        265  ATKAIGKKI-GNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKN-EELKEKIDTAKQCS 322

Qy        302  EDFTKKLEGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLAKAAKEML 361
               :|| ||| ||||||:||:|||  :|| |||  ||||||||||||||||||||:||::
Db        323  TEFTNKLKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTL 382

Qy        362  ANSVKEL 368
               |:|||||
Db        383  KNAVKEL 389


RESULT 13
US-09-596-746A-26
Sequence 26, Application US/0959746A
GENERAL INFORMATION:
;   APPLICANT: Dattwyler, Raymond J.
;   APPLICANT: Seinost, Gerald
;   APPLICANT: Dykhuizen, Daniel
;   APPLICANT: Luft, Benjamin J.
;   APPLICANT: Maria J.C. Gomes-Solecki
;   TITLE OF INVENTION: Groups of Borrelia burgdorferi and
;   TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
;   FILE REFERENCE: 2631.1002-001
;   CURRENT APPLICATION NUMBER: US-09/596,746A
;   CURRENT FILING DATE: 2000-06-19
;   PRIOR APPLICATION NUMBER: US 60/140,042
;   PRIOR FILING DATE: 1999-06-18
;   NUMBER OF SEQ ID NOS: 84
;   SOFTWARE: FastSEQ for Windows Version 4.0
;   SEQ ID NO 26
;   LENGTH: 374
;   TYPE: PRT
;   FEATURE:
;   OTHER INFORMATION: Artificial Sequence
;   ORGANISM: Ospc Chimera
US-09-596-746A-26

Query Match             82.8%;  Score 1489;  DB 19;  Length 374;
Best Local Similarity   83.4%;  Pred. No. 3.4e-101;
Matches   312;  Conservative   22;  Mismatches   34;  Indels    6;  Gaps    2;

Qy          1  MACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKA  60
               |||||||||||:||||||||||||||||||||||||||||||||||||||:|||::|||
Db          1  MACNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLSSIDEIAAKA  60

Qy         61  IGKKI-GNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTN 119
               |||||   ||||:  |  :|| |||||||  |:  |  |||||||||||: |  |||||
Db         61  IGKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKNEELKEKIDAAKKCSEFTN 120

Qy        120  KLKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVK 179
               ||| | |:  |  :|| ::|  |||  |  ||:| ||:|||| :::|||||||:|:||:||
Db        121  KLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 180

Qy        180  ELTSPIVHG-----NNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETL 234
               |||||:||      |||||||||||||||||||||||||||||||||||||||||||||||
Db        181  ELTSPVVAESPAMWNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETL 240

Qy        235  LASIDELATKAIGKKIQQNGGLAVEAGHNGTLLAGAYTISKLLTQKLDGLKNSEKLKEKI 294
               ||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||
Db        241  LASIDELATKAIGKKIQONGGLAVEAGHNGTLLAGAYTISKLLTQKLDGLKNSEKLKEKI 300

Qy        295  ENAKKCSEDFTKKLEGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLA 354
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301  ENAKKCSEDFTKKLEGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLA 360
```

Qy   355 KAAKEMLANSVKEL 368
          ||||||||||||||
Db   361 KAAKEMLANSVKEL 374

RESULT 14
US-09-596-746-30
Sequence 30, Application US/0959967746
GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Selnost, Gerald
; APPLICANT: Dykhuizen, Danial
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746
; CURRENT FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 377
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: OspC Chimera
US-09-596-746-30

Query Match   82.6%;   Score 1485.5; DB 19;   Length 377;
Best Local Similarity   82.5%;   Pred. No. 6.2e-101;
Matches 311; Conservative 23; Mismatches 34; Indels 9; Gaps 2;

Qy   1 MACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKA 60
          |||||||||||||:|||||||||||||||||||||||||||||| ||:|:||||||:| ||
Db   1 MACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKA 60

Qy   61 IGKKI-GNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTN 119
          ||||| ||||:|||||||||||:|||:|:|||||||||||||:|||||:|||||||||| |||
Db   61 IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN 120

Qy   120 KLKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVK 179
          ||| :| :|| :|:|| | :|| :|| || | |||||||||||||::|: || ||:| |
Db   121 KLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 180

Qy   180 ELTSPIVHG--------NNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEI 231
          |||:|:|:| |||||||||||||||||||||||||||||||||||||||||||
Db   181 ELTSPVVAESPKKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEI 240

Qy   232 ETLLASIDELATKAIGKKIQQNGGLAVEAGHNGTLLAGAYTISKLIITQKLDGLKNSEKLK 291
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 ETLLASIDELATKAIGKKIQQNGGLAVEAGHNGTLLAGAYTISKLIITQKLDGLKNSEKLK 300

Qy   292 EKIENAKKCSEDFTKKLEGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVE 351
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 EKIENAKKCSEDFTKKLEGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVE 360

Qy   352 NLAKAAKEMLANSVKEL 368
          ||||||||||||||||
Db   361 NLAKAAKEMLANSVKEL 377

RESULT 15
US-09-596-746A-30
Sequence 30, Application US/0959967746A
GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Selnost, Gerald
; APPLICANT: Dykhuizen, Danial
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746A
; CURRENT FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 377
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: OspC Chimera
US-09-596-746A-30

Query Match   82.6%;   Score 1485.5; DB 19;   Length 377;
Best Local Similarity   82.5%;   Pred. No. 6.2e-101;
Matches 311; Conservative 23; Mismatches 34; Indels 9; Gaps 2;

Qy   1 MACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKA 60
          |||||||||||||:|||||||||||||||||||||||||||||| ||:|:||||||:| ||
Db   1 MACNNSGKDGNTSANSADESVKGPNLTEINKKITDSNAVLLAVKEVEALLSSIDEIAAKA 60

Qy   61 IGKKI-GNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTN 119
          ||||| ||||:|||||||||||:|||:|:|||||||||||||:|||||:|||||||||| |||
Db   61 IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN 120

Qy   120 KLKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVK 179
          ||| :| :|| :|:|| | :|| :|| || | |||||||||||||::|: || ||:| |
Db   121 KLKEKHTDLGKEGVTDADAKEAILKANGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 180

Qy   180 ELTSPIVHG--------NNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEI 231
          |||:|:|:| |||||||||||||||||||||||||||||||||||||||||||
Db   181 ELTSPVVAESPKKPSMVNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEI 240

Qy   232 ETLLASIDELATKAIGKKIQQNGGLAVEAGHNGTLLAGAYTISKLIITQKLDGLKNSEKLK 291
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 ETLLASIDELATKAIGKKIQQNGGLAVEAGHNGTLLAGAYTISKLIITQKLDGLKNSEKLK 300

Qy   292 EKIENAKKCSEDFTKKLEGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVE 351
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 EKIENAKKCSEDFTKKLEGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVE 360

Qy   352 NLAKAAKEMLANSVKEL 368
          ||||||||||||||||
Db   361 NLAKAAKEMLANSVKEL 377

Search completed: March 18, 2002, 10:08:46
Job time: 974 sec

OM protein - protein search, using sw model

```
                    GenCore version 4.5
                    Copyright (c) 1993 - 2000 Compugen Ltd.

Run on:             March 18, 2002, 09:55:35 ; Search time 55.5 Seconds
                                               (without alignments)
                                               149.211 Million cell updates/sec

Title:              1798
Perfect score:      1798
Sequence:           1 MACNNSGKDGNASANSADES..........AVENLAKAAKEMLANSVKEL 368

Scoring table:      BLOSUM62
                    Gapop 10.0 , Gapext 0.5

Searched:           US-09-596-746A-42
                    212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters:   212252

Maximum DB seq length: 2000000000
Minimum DB seq length: 0

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Issued_Patents_AA:*
             1:  /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*
             2:  /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*
             3:  /cgn2_6/ptodata/2/1aa/6A_COMB.pep:*
             4:  /cgn2_6/ptodata/2/1aa/6B_COMB.pep:*
             5:  /cgn2_6/ptodata/2/1aa/PCTUS_COMB.pep:*
             6:  /cgn2_6/ptodata/2/1aa/backfiles1.pep:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Match Length	DB	ID	Description
1	699	38.9	588	4	US-08-235-836C-122	Sequence 122, App
2	698.5	38.9	466	4	US-08-235-836C-107	Sequence 107, App
3	665	37.0	466	4	US-08-235-836C-110	Sequence 110, App
4	657.5	36.6	212	1	US-08-158-353-2	Sequence 2, Appl
5	644.5	35.8	209	4	US-09-196-293-15	Sequence 15, Appl
6	644.5	35.8	210	1	US-08-158-353-3	Sequence 3, Appl
7	639.5	35.6	210	4	US-08-235-836C-30	Sequence 30, Appl
8	637.5	35.5	210	4	US-08-209-603E-15	Sequence 15, Appl
9	600.5	33.4	212	1	US-08-031-295-2	Sequence 2, Appl
10	600.5	33.4	212	4	US-07-903-580-2	Sequence 2, Appl
11	588.5	32.7	212	4	US-08-158-353-4	Sequence 4, Appl
12	583.5	32.5	212	4	US-08-235-836C-34	Sequence 34, Appl
13	578.5	32.2	212	4	US-09-196-293-11	Sequence 11, Appl
14	578.5	32.2	212	4	US-08-209-603E-11	Sequence 11, Appl
15	563	31.3	207	4	US-08-235-836C-36	Sequence 36, Appl
16	553	30.8	209	4	US-08-235-836C-32	Sequence 32, Appl
17	168.5	9.4	1196	2	US-08-144-121-4	Sequence 4, Appl
18	168.5	9.4	1196	2	US-08-735-893-4	Sequence 4, Appl
19	152.5	8.5	630	4	US-08-973-462-9	Sequence 9, Appl
20	151.5	8.4	1181	4	US-08-053-614-4	Sequence 4, Appl
21	151.5	8.4	1181	1	US-08-316-397B-4	Sequence 4, Appl
22	151.5	8.4	1181	2	US-09-034-306-4	Sequence 4, Appl
23	151.5	8.4	1181	4	US-09-259-437-4	Sequence 4, Appl
24	151.5	8.4	1181	5	PCT-US93-09782-4	Sequence 4, Appl
25	151.5	8.4	2482	1	US-08-328-254-6	Sequence 6, Appl
26	150.5	8.4	3248	1	US-08-353-700-1	Sequence 1, Appl
27	150.5	8.4	3248	5	PCT-US95-16216-1	Sequence 1, Appl
28	146.5	8.1	1164	4	US-08-923-992A-2	Sequence 2, Appl1
29	146.5	8.1	1786	4	US-08-973-462-8	Sequence 8, Appl1
30	145	8.1	1147	3	US-08-470-260-5	Sequence 5, Appl1
31	145	8.1	1147	3	US-08-471-491-5	Sequence 5, Appl1
32	145	8.1	1147	4	US-08-466-662-5	Sequence 5, Appl1
33	145	8.1	3289	2	US-08-477-451-2	Sequence 2, Appl1
34	142	7.9	688	3	US-09-141-047-8	Sequence 8, Appl1
35	141.5	7.9	2101	4	US-08-460-309-4	Sequence 4, Appl1
36	141.5	7.9	3111	2	US-08-125-077-4	Sequence 4, Appl1
37	141	7.8	3111	2	US-08-466-390-4	Sequence 4, Appl1
38	141	7.8	2101	1	US-08-470-950-4	Sequence 4, Appl1
39	141	7.8	2101	1	US-08-467-781-4	Sequence 4, Appl1
40	141	7.8	2101	1	US-08-195-487-4	Sequence 4, Appl1
41	141	7.8	2101	4	US-08-483-924-4	Sequence 4, Appl1
42	141	7.8	2101	4	US-09-452-294-1	Sequence 1, Appl1
43	141	7.8	2101	5	PCT-US93-06160-4	Sequence 4, Appl1
44	138.5	7.7	1128	4	US-08-923-992A-6	Sequence 6, Appl1
45	136.5	7.6	679	4	US-08-913-942-15	Sequence 15, Appl1

ALIGNMENTS

```
RESULT 1
US-08-235-836C-122
Sequence 122, Application US/08258336C
Patent No. 6248562
GENERAL INFORMATION:
    APPLICANT: Dunn, John J.
    APPLICANT: Luft, Benjamin J.
    TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising
    TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor
    NUMBER OF SEQUENCES: 144
    CORRESPONDENCE ADDRESS:
        ADDRESSEE: Brookhaven National Laboratory.
        STREET:
        CITY: Upton
        STATE: NY
        COUNTRY: USA
        ZIP: 11973
    COMPUTER READABLE FORM:
        MEDIUM TYPE: Floppy disk
        COMPUTER: IBM PC compatible
        OPERATING SYSTEM: PC-DOS/MS-DOS
        SOFTWARE: Patentin Release #1.0, Version #1.25
    CURRENT APPLICATION DATA:
        APPLICATION NUMBER: US/08/235,836C
        FILING DATE: 29-APR-1994
        CLASSIFICATION: 435
    PRIOR APPLICATION DATA:
        APPLICATION NUMBER: US 08/148,191
        FILING DATE: 01-11-93
    ATTORNEY/AGENT INFORMATION:
        NAME: Bogosian, Margaret C.
        REGISTRATION NUMBER: 25,324
        REFERENCE/DOCKET NUMBER: BNL93-28A
    TELECOMMUNICATION INFORMATION:
        TELEPHONE: (516) 282-7338
        TELEFAX: (516) 282-3729
INFORMATION FOR SEQ ID NO: 122:
    SEQUENCE CHARACTERISTICS:
        LENGTH: 588 amino acids
        TYPE: amino acid
        TOPOLOGY: linear
        MOLECULE TYPE: protein
US-08-235-836C-122

Query Match 38.9%;  Score 699;  DB 4;  Length 588;
Best Local Similarity 41.5%;
Matches 195;  Conservative 49;  Mismatches 104;  Indels 122;  Gaps 17;
Pred. No. 5.1e-45;

Qy    7 GKDGNASANSADESVKG-----------PNLTEISKKITESNAVVLAVKEVET---LLASID 54
```

RESULT 2

US-08-235-836C-107
Sequence 107, Application US/08235836C
Patent No. 6248562

GENERAL INFORMATION:
APPLICANT: Dunn, John J.
APPLICANT: Luft, Benjamin J.
TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising
TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor
NUMBER OF SEQUENCES: 144
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brookhaven National Laboratory
STREET:
CITY: Upton
STATE: NY
COUNTRY: USA
ZIP: 11973
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/235,836C
FILING DATE: 29-APR-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/148,191
FILING DATE: 01-11-93
ATTORNEY/AGENT INFORMATION:
NAME: Bogoslan, Margaret C.
REGISTRATION NUMBER: 25,324
REFERENCE/DOCKET NUMBER: BNL93-28A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 282-7338
TELEFAX: (516) 282-3729

INFORMATION FOR SEQ ID NO: 107:
SEQUENCE CHARACTERISTICS:
LENGTH: 466 amino acids
TYPE: amino acid
TOPOLOGY: linear

```
Db    9 GSQKENDLNLEDSSKKSHQNAKQDLPAVTEDSVSLFNGNKIFVSKEKNSSGKYDLRATID 68

Qy   55 ELATKAIGKKIGNNG---LEANQ-SKNTSLLSGAYAISDIIAEKLNVLKNEELKEKIDTA 110

Db   69 QVELKGTSDK--NNSGTLEGSKPDKSKVKLITVSADLNTVTLEAFDA-SNQKISSKV-TK 124

Qy  111 KQCS-TEFT---NKL----KSEHAVIGLDNLTD-DNAQRAI--------------- 142

Db  125 KQGSTEEETLKANKLDSKKLTRSNGTTLEYSQITDADNATKAVETLKNSIKLEGSLVGK 184

Qy  143 --------LKKHANKD------ 150

Db  185 TTVEIKEGTVTLIKRELEKDGKVKVFLNDTAGSNKKTGKWEDSTSTLTISADSKKTKDLVF 244

Qy  151 -----KGAAELEKLFKAVENLSKAAQDTLKNAVKELTSPIVHGNNSGKDGNTS 198

Db  245 LITDGTITVQQYNTAGTSLEGSASEIKNLSE-----LKNALK--GHPM--GNNSGKDGNTS 295

Qy  199 ANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDELATKAIGKKIQONGGLAV 258

Db  296 ANSADESVKGPNLTEISKKIITDSNAVLLAVKEVEALLSSIDEIAAKAIGKKIHQNNGLDT 355

Qy  259 EAGHNGTLLAGAYTISKLITQKLIDGLKNSEKLKEKIENAKKCSEDTKKLEGEHAQLGIE 318

Db  356 EYNHNGSLLAGRYAISTLIKQKLIDGLKN-EGLKEKIDAAKKCSETFTNKLKEKHTDLGKE 414

Qy  319 NVTDENAKKAILITDAAKDKGAAELEKLFKAVENLAKAAKEMLANSVKEL 368

Db  415 GVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKEL 464
```

MOLECULE TYPE: protein
US-08-235-836C-107

```
Query Match    Best Local Similarity   38.8%;  Score 698.5; DB 4; Length 466;
               47.7%; Pred. No. 4.1e-45;
Matches  177;  Conservative 43; Mismatches 114; Indels 37; Gaps 8;
```

```
Qy   19 ESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKAIGKKIGNNGLEANQS-KN 77

Db  100 EVFKEDGKTLVSKKVTSKDKSSTEEKFENEKGEVS-EKIITRADGTRLEYTGIKSDGSGKA 158

Qy   78 TSLLSGAYAISDLIAEKLNVLKNE---ELKEKIDTAKQCSTEFTNKLKSEHAVIGLDNLT 134

Db  159 KEVLKGYVLEGTLTAEKTLVVKEGTVTLISKNISKSGEVSVELNDTDSS--AATKKTAAW 216

Qy  135 DDNAQRAILKKHANKDK--------GAAELEKLFKAVENLSKAAQDTLKNA 177

Db  217 NDSTSTLTISADSKKTKDLVFLITDGTITVQQYNTAGTSLEGSASEIKNLSE-----LKNA 271

Qy  178 VKELTSPIVHGNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLAS 237

Db  272 LK-------MANNSGKDGNTSANSADESVKGPNLTEISKKIITDSNAVLLAVKEVEALLSS 324

Qy  238 IDELATKAIGKKIQONGGLAVEAGHNGTLLAGAYTISKLITQKLIDGLKNSEKLKEKIENA 297

Db  325 IDEIAAKAIGKKIHQNNGLDTEYNHNGSLLAGRYAISTLIKQKLIDGLKN-EGLKEKIDAA 383

Qy  298 KKCSEDTKKLEGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLAKAA 357

Db  384 KKCSETFTNKLKRAKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAA 443

Qy  358 KEMLANSVKEL 368

Db  444 KEMLANSVKEL 454
```

RESULT 3

US-08-235-836C-110
Sequence 110, Application US/08235836C
Patent No. 6248562

GENERAL INFORMATION:
APPLICANT: Dunn, John J.
APPLICANT: Luft, Benjamin J.
TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising
TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor
NUMBER OF SEQUENCES: 144
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brookhaven National Laboratory
STREET:
CITY: Upton
STATE: NY
COUNTRY: USA
ZIP: 11973
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US 08/235,836C
FILING DATE: 29-APR-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/148,191
FILING DATE: 01-11-93
ATTORNEY/AGENT INFORMATION:
NAME: Bogoslan, Margaret C.
REGISTRATION NUMBER: 25,324
REFERENCE/DOCKET NUMBER: BNL93-28A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 282-7338
TELEFAX: (516) 282-3729

INFORMATION FOR SEQ ID NO: 110:
: SEQUENCE CHARACTERISTICS:
:   LENGTH: 466 amino acids
:   TYPE: amino acid
:   TOPOLOGY: linear
:   MOLECULE TYPE: protein
US-08-235-836C-110

Query Match   37.0%; Score 665; DB 4; Length 466;
Best Local Similarity   46.9%; Pred. No. 1.3e-42;
Matches 165; Conservative 47; Mismatches 104; Indels 36; Gaps 8;

Qy   1 MACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKVEVETLLASIDELATKA 60
         ::|||||||| |||||||||||||||||||||||||||||||||||||:||:||:|:| ||
Db   17 ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA 76

Qy   61 IGKKI-GNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTN 119
         |:||: |||||||||||:|||||||||||||||||:||||||:||||||||||||||||
Db   77 IGKKIHQNNGLDTEYNHNGSLLAGRYAISTLIKQKLCDGLIKNEGLKEKIDAAKKCSEFTN 136

Qy 120 KLKSEHAVLGLDNLTLDDNAQRAILKKHANKDKGAAAELEKLFKAVENLSKAAQDTLKNAVK 179
         ||:|||||| |||:||||||||||||||||||||||||||||||||||||||:||||||
Db 137 KLKAKHFDLGKEGVTDADAKEALLKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 196

Qy 180 ELTSPIVHGNNSGKDGNTSA----NSADESVKGPNLTEISKKITESNAVVLAVKEIEFLL 235
         |||||||||:|:| ||||||||||||||||||||||||:||||||
Db 197 ELTSPVVAESPKPKQNVSSLDEKNSVSVDLPGEMKVLVSKEKNKDGKY------DLI 248

Qy 236 ASIDELATKAIGKKIQQNGGL-AVEAGHNGTLIA--------------GAYTISKLIT 278
         ||||||||:|:|||||| ||:||||||:
Db 249 ATVDKLELKGTSDKDNNGSGVLEGVKADKSKVKLTISDDLGQTTLEVFKEDGKTLVSKKVT 308

Qy 279 QKLDGLKNSEKLKEKIENAKKCSEDFTKKLEGEHAQ-LGIENVTDENAKKAI 329
         |:| || || ||||:|||| ||:|||||||||||
Db 309 SK-DKSSTEEKENEKGE----VSEKIITRADGTRLEYTGIKSDGSGKAKEVL 355

RESULT   4
US-08-158-353-2
: Sequence 2, Application US/08158353
: Patent No. 5620062
GENERAL INFORMATION:
  APPLICANT: Padula, Steven J.
  TITLE OF INVENTION: Methods for Diagnosing Early Lyme
  TITLE OF INVENTION: Disease
  NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
  ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
  STREET: Two Militia Drive
  CITY: Lexington
  STATE: MA
  COUNTRY: USA
  ZIP: 02173
COMPUTER READABLE FORM:
  MEDIUM TYPE: Floppy disk
  COMPUTER: IBM PC compatible
  OPERATING SYSTEM: PC-DOS/MS-DOS
  SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/158,353
  FILING DATE:
  CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
  NAME: Carroll, Alice O.
  REGISTRATION NUMBER: 33,542
  REFERENCE/DOCKET NUMBER: UCT93-05
TELECOMMUNICATION INFORMATION:
  TELEPHONE: 617-861-6240
  TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
:   LENGTH: 212 amino acids

: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-08-158-353-2

Query Match   36.6%; Score 657.5; DB 1; Length 212;
Best Local Similarity   72.9%; Pred. No. 1.7e-42;
Matches 140; Conservative 20; Mismatches 29; Indels 3; Gaps 2;

Qy   1 MACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKVEVETLLASIDELATKA 60
         ::|||||||| |||||||||||||||||||||||||||||||||||||:||:||:|:| ||
Db   17 ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEVAKKA 76

Qy   61 IGKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKN-EELKEKIDTAKQCSTEFTN 119
         ||||||||||||||||||||||||||||||||||||:|| ||||||||||||:||||||
Db   77 IGNLIAQNGLNAGANQNGSLLAGAVVISTLIAEKLDGLIKNSEELKEKIDAAKKCNAFTD 136

Qy 120 KLKSEHAVLGLDN--LTLDDNAQRAILKKHANKDKGAAAELEKLFKAVENLSKAAQDTLKNA 177
         |||| ||||||| ||||||||||||||| ||||||||||||||||||||||||||||
Db 137 KLKSSHAELGIANGAATDANAKAAILKTNGTKDKGAQOLEKLFESVKNLSKAAQETLNNS 196

Qy 178 VKELTSPIVHGN 189
         |||||:||:||
Db 197 VKELTSPVVAEN 208

RESULT   5
US-09-196-293-15
: Sequence 15, Application US/09196293
: Patent No. 6183755
GENERAL INFORMATION:
  APPLICANT: Fuchs, Renate
  APPLICANT: Motz, Manfred
  APPLICANT: Soutscheck, Erwin
  APPLICANT: Wilske, Bettina
  APPLICANT: Preac-Mursic, Vera
  TITLE OF INVENTION: Active proteins from Borrelia
  TITLE OF INVENTION: burgdorferi
  FILE REFERENCE: 738.001US2
  CURRENT APPLICATION NUMBER: US/09/196,293
  CURRENT FILING DATE: 1998-11-19
  EARLIER APPLICATION NUMBER: US 08/209,603
  EARLIER FILING DATE: 1994-03-10
  EARLIER APPLICATION NUMBER: US 07/862,535
  EARLIER FILING DATE: 1992-06-19
  EARLIER APPLICATION NUMBER: WO PCT/EP90/02282
  EARLIER FILING DATE: 1990-12-21
  EARLIER APPLICATION NUMBER: DE P39 42 728.5
  EARLIER FILING DATE: 1989-12-22
  EARLIER APPLICATION NUMBER: DE P40 18 988.0
  EARLIER FILING DATE: 1990-06-13
  NUMBER OF SEQ ID NOS: 16
  SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 15
: LENGTH: 209
: TYPE: PRT
: ORGANISM: Borrelia burgdorferi
US-09-196-293-15

Query Match   35.8%; Score 644.5; DB 4; Length 209;
Best Local Similarity   75.6%; Pred. No. 1.6e-41;
Matches 136; Conservative 15; Mismatches 28; Indels 1; Gaps 1;

Qy 189 NNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKVEIETLLASIDELATKAIGK 248
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db   20 NNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGK 79

Qy 249 KIQQNGGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKEKIENAKKCSEDFTKKL 308
         ||| || || || | ||||| | | ||| | ||| |||||| ||||||||||||||||
Db   80 KIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLCDGLKN-EGLKEKIDAAKKCSEFTNKL 138

Qy   309 EGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLAKAAKEMLANSVKEL 368
          :|| |||| ||| ||||| |:||||||||| |:|:|||||||||||||||||||||||
Db   139 KEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKEL 198

RESULT 6
US-08-158-353-3
Sequence 3, Application US/0815353
Patent No. 5620862
GENERAL INFORMATION:
  APPLICANT: Padula, Steven J.
  TITLE OF INVENTION: Methods for Diagnosing Early Lyme
  TITLE OF INVENTION: Disease
  NUMBER OF SEQUENCES: 7
  CORRESPONDENCE ADDRESS:
    ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
    STREET: Two Militia Drive
    CITY: Lexington
    STATE: MA
    COUNTRY: USA
    ZIP: 02173
  COMPUTER READABLE FORM:
    MEDIUM TYPE: Floppy disk
    COMPUTER: IBM PC compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: PatentIn Release #1.0, Version #1.25
  CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/08/158,353
    FILING DATE:
  CLASSIFICATION: 435
  ATTORNEY/AGENT INFORMATION:
    NAME: Carroll, Alice O.
    REGISTRATION NUMBER: 33,542
    REFERENCE/DOCKET NUMBER: UCT93-05
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: 617-861-6240
    TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 3:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 210 amino acids
    TYPE: amino acid
    STRANDEDNESS: single
    TOPOLOGY: linear
  MOLECULE TYPE: protein
US-08-158-353-3

Query Match   35.8%; Score 644.5; DB 1; Length 210;
Best Local Similarity   75.6%; Pred. No. 1.6e-41;
Matches 136; Conservative 15; Mismatches 28; Indels   1; Gaps   1;

Qy   189 NNSGKKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDELATKAIGK 248
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   20 NNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGK 79

Qy   249 KIQQNGGLAVEAGHNGTLLAGAYTISKLITQKLIDGLKNSEKLKEKIENAKKCSEDFTKKL 308
          :||| ||| |||||| |||||||||||| |||||||||||| ||||||||| |||||||
Db   80 KIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLIDGLKN-EGLKEKIDAAKKCSETFTNKL 138

Qy   309 EGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLAKAAKEMLANSVKEL 368
          :|| |||| ||| ||||| |:||||||||| |:|:|||||||||||||||||||||||
Db   139 KEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKEL 198

RESULT 7
US-08-235-836C-30
Sequence 30, Application US/08235836C
Patent No. 6248562
GENERAL INFORMATION:
  APPLICANT: Dunn, John J.
  APPLICANT: Luft, Benjamin J.
  TITLE OF INVENTION: No. 6248562e1 Chimeric Proteins Comprising

---

TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor
NUMBER OF SEQUENCES: 144
CORRESPONDENCE ADDRESS:
  ADDRESSEE: Brookhaven National Laboratory
  STREET:
  CITY: Upton
  STATE: NY
  COUNTRY: USA
  ZIP: 11973
COMPUTER READABLE FORM:
  MEDIUM TYPE: Floppy disk
  COMPUTER: IBM PC compatible
  OPERATING SYSTEM: PC-DOS/MS-DOS
  SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/235,836C
  FILING DATE: 29-APR-1994
  CLASSIFICATION: 435
PRIOR APPLICATION DATA:
  FILING DATE: 01-11-93
  APPLICATION NUMBER: US 08/148,191
ATTORNEY/AGENT INFORMATION:
  NAME: Boggosian, Margaret C.
  REGISTRATION NUMBER: 25,324
  REFERENCE/DOCKET NUMBER: BNL93-28A
TELECOMMUNICATION INFORMATION:
  TELEPHONE: (516) 282-7338
  TELEFAX: (516) 282-3729
INFORMATION FOR SEQ ID NO: 30:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 210 amino acids
    TYPE: amino acid
    TOPOLOGY: linear
  MOLECULE TYPE: protein
US-08-235-836C-30

Query Match   35.6%; Score 639.5; DB 4; Length 210;
Best Local Similarity   75.0%; Pred. No. 3.9e-41;
Matches 135; Conservative 15; Mismatches 29; Indels   1; Gaps   1;

Qy   189 NNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDELATKAIGK 248
          ||||||||||||||||||||||||||||||||::||||||||||||||||::|||||
Db   20 NNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGK 79

Qy   249 KIQQNGGLAVEAGHNGTLLAGAYTISKLITQKLIDGLKNSEKLKEKIENAKKCSEDFTKKL 308
          :||| ||| |||||| |||||||||||| |||||||||||| ||||||||| |||||||
Db   80 KIHQNNGLDTEYNHNGSLLAGRYAISTLIKQKLIDGLKN-EGLKEKIDAAKKCSETFTNKL 138

Qy   309 EGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLAKAAKEMLANSVKEL 368
          :|| |||| ||| |||||| |:||||||||| |:|:|||||||||||||||||||||||
Db   139 KEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKEL 198

RESULT 8
US-08-209-603E-15
Sequence 15, Application US/08209603E
Patent No. 6248538
GENERAL INFORMATION:
  APPLICANT: FUCHS, RENATE
  APPLICANT: WILSKE, BETTINA
  APPLICANT: PREAC-MURSIC, VERA
  APPLICANT: MOTZ, MANFRED
  APPLICANT: SOUTSCHEK, ERWIN
  TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PROTEINS
  TITLE OF INVENTION: FROM BORRELIA BURGDORFERI
  NUMBER OF SEQUENCES: 15
  CORRESPONDENCE ADDRESS:
    ADDRESSEE: BROOKS HAIDT HAFFNER & DELAHUNTY
    STREET: 99 PARK AVENUE
    CITY: NEW YORK
    STATE: NY
    COUNTRY: USA

ZIP: 10016
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" FLOPPY DISC
COMPUTER: AT&T - IBM COMPATIBLE
OPERATING SYSTEM: MS-DOS Version 6.2
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/209,603E
FILING DATE: 10-MAR-1994
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP90/02282
FILING DATE: 21-DEC-1990
APPLICATION NUMBER: US 07/862,535
FILING DATE: 19-JUN-1992
ATTORNEY/AGENT INFORMATION:
NAME: ROBINSON, WILLIAM R.
REGISTRATION NUMBER: 27,224
REFERENCE/DOCKET NUMBER: LKR-9217-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 697-3355
TELEFAX: (212) 557-5635
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 210
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
DESCRIPTION: PROTEIN
HYPOTHETICAL: N/A
ANTI-SENSE: N/A
FRAGMENT TYPE: N/A
ORIGINAL SOURCE:
ORGANISM: B. BURGDORFERI
IMMEDIATE SOURCE:
LIBRARY: DSM 5662
POSITION IN GENOME: N/A
FEATURE:
IDENTIFICATION METHOD: amino acid analysis
PUBLICATION INFORMATION: N/A
US-08-209-603E-15

```
Query Match     35.5%;  Score 637.5;  DB 4;  Length 210;
Best Local Similarity  75.0%;  Pred. No. 5.5e-41;
Matches  135;  Conservative  15;  Mismatches  29;  Indels  1;  Gaps  1;

Qy   189  NNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDELATKAIGK  248
          ||||||||||||||||||||||:||||||:.|||.|||||:||||:|:||.|:|:|||||
Db    20  NNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVLVEALLSSIDEIAAKAIGK   79

Qy   249  KIQQNGGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKEKIENAKKCSEDFIKKL  308
          |:||||  | ||||||:|||||||||||||||||||||||||||||||||||||||||||
Db    80  KIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKN-EGLKEKIDAAKKCSETFTNKL  138

Qy   309  EGEHAQLGIENVTDENAKKAILITDAARKDKGAAELEKLFKAVENLAKAAKEMLANSVKEL  368
          :||:|| | ||| |||:|:.|||| |:|:|||||||:|||||||||||||||||||||||
Db   139  KEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKEL  198
```

RESULT 9
US-08-031-295-2
Sequence 2, Application US/0803129S
Patent No. 5530103
GENERAL INFORMATION:
APPLICANT: LIVEY, Ian
APPLICANT: DORNER, Freidrich
TITLE OF INVENTION: METHOD AND COMPOSITION FOR THE
TITLE OF INVENTION: PREVENTION OF LYME DISEASE
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/031,295
FILING DATE: 19930312
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/903,580
FILING DATE: 25-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/824,161
FILING DATE: 22-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/727,245
FILING DATE: 11-JUL-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/142 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 212 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-031-295-2

```
Query Match     33.4%;  Score 600.5;  DB 1;  Length 212;
Best Local Similarity  70.2%;  Pred. No. 3.4e-38;
Matches  127;  Conservative  20;  Mismatches  33;  Indels  1;  Gaps  1;

Qy   189  NNSGKDGNT-SANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDELATKAIG  247
          ||||||  :| |||.||||.|||||||||||||||:|||.:|||.|.|:||||||||||||
Db    20  NNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETLVSSIDELATKAIG   79

Qy   248  KKIQQNGGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKEKIENAKKCSEDFTKK  307
          ||||||| || || ||||||||||||||:|||||||||||||:||||||||||||||||
Db    80  KKIQQNNGLGANADKNGSLLAGAYAISTLITEKLKALKNSGELKAKIEDAKKCSEDFTKK  139

Qy   308  LEGEHAQLGIENVTDENAKKAILITDAARKDKGAAELEKLFKAVENLAKAAKEMLANSVKE  367
          |||||||:||: ||| |||:|||| |:|:|||||||:|||||||:.|:|||:|||||||||
Db   140  LAAGHAQLGIDGATDNDSKEAILKTNGTKTKGAEELVKLSESVASLSKAAQEASANSVKE  199

Qy   368  L  368
          |
Db   200  L  200
```

RESULT 10
US-07-903-580-2
Sequence 2, Application US/07903580
Patent No. 6221363
GENERAL INFORMATION:
APPLICANT: LIVEY, Ian
APPLICANT: DORNER, Freidrich
TITLE OF INVENTION: METHOD AND COMPOSITION FOR THE
TITLE OF INVENTION: PREVENTION OF LYME DISEASE
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria

STATE: Virginia
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/903,580
FILING DATE: 1920625
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/824,161
FILING DATE: 22-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/727,245
FILING DATE: 11-JUL-1991
ATTORNEY/AGENT INFORMATION:
NAME: Bent, Stephen A.
REGISTRATION NUMBER: 29,768
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
REFERENCE/DOCKET NUMBER: 30472/131 IMMU
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 212 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-903-580-2

Query Match
Best Local Similarity   33.4%;   Score 600.5; DB 4; Length 212;
Matches 127; Conservative 20; Mismatches 33;   Indels 1; Gaps 1;

```
Qy  189 NNSGKDGNT-SANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDELATKAIG 247
            |||||||| :||||||||||||||||||||| ||||||||| ||||||||||||||
Db   20 NNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETLVSSIDELATKAIG 79

Qy  248 KKIQONGGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKEKIENAKKCSEDFTKK 307
            ||||| || |  ||||||||||| ||:|||  | || || |  ||  |||||||||||
Db   80 KKIQONNGLGANADKNGSLLAGAYAISTLITEKLRALKNSGELKAKIEDAKKCSEDFTKK 139

Qy  308 LEGEHAQLGIENVTDENAKRAILITDAAKDKGGAAELEKLFKKAVENLAKAAKEMLANSVKE 367
            | |||||||  |  :||||| |:   |  || || |||| :| ||| | || ||||
Db  140 LAAGHAQLGIDGATDNDSKEAILKTNGTKTKGAEELVKLSESVASLSKAAQEASANSVKE 199
```

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/158,353
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Carroll, Alice O.
REGISTRATION NUMBER: 33,542
REFERENCE/DOCKET NUMBER: UCT93-05
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-861-6240
TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 212 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-158-353-4

Query Match
Best Local Similarity   32.7%;   Score 588.5; DB 1; Length 212;
Matches 130; Conservative 20; Mismatches 36;   Indels 3; Gaps 3;

```
Qy    1 MACNNSGKDG-NASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATK 59
           ::|||||| | |||||||||||||||||||||||||:|||||||||||||||||||||
Db   17 ISCNNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETLVSSIDELATKAKK 76

Qy   60 AIGKKI-GNNGLEANQSKNTSLLSGAYAISDIIAEKLNVLKN-EELKEKIDTAKQCSTEF 117
           ||:||| | ||| ||  |  ||| ||||:|| |:| ||| |: :||||| :|:|:|||
Db   77 AIGQKIDNNNGLAALNNQNGSLLAGAYAISTLITEKLSKLKNLEELKTEIAKAKKCSEEF 136

Qy  118 TNKLKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNA 177
           |||||| || || :| |||||| |||| ||:|:||| |||||||| ||||::|| |||
Db  137 TNKLKSGHADLGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGLLKAAQVALTNS 196

Qy  178 VKELTSPIV 186
           ||||||:|
Db  197 VKELTSPVV 205
```

RESULT 11
US-08-158-353-4
Sequence 4, Application US/08158353
Patent No. 5620862
GENERAL INFORMATION:
APPLICANT: Padula, Steven J.
TITLE OF INVENTION: Methods for Diagnosing Early Lyme Disease
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Militia Drive
CITY: Lexington
STATE: MA
COUNTRY: USA
ZIP: 02173
COMPUTER READABLE FORM:

RESULT 12
US-08-235-836C-34
Sequence 34, Application US/08235836C
Patent No. 6248562
GENERAL INFORMATION:
APPLICANT: Dunn, John J.
APPLICANT: Luft, Benjamin J.
TITLE OF INVENTION: No. 6248562e1 Chimeric Proteins Comprising Borrelia Polypeptides and Uses Therefor
NUMBER OF SEQUENCES: 144
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brookhaven National Laboratory
STREET:
CITY: Upton
STATE: NY
COUNTRY: USA
ZIP: 11973
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/235,836C
FILING DATE: 29-APR-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/148,191
FILING DATE: 01-11-93
ATTORNEY/AGENT INFORMATION:
NAME: Bogosian, Margaret C.
REGISTRATION NUMBER: 25,324
REFERENCE/DOCKET NUMBER: BNL93-28A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 282-7338
TELEFAX: (516) 282-3729
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 212 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-235-836C-34

Query Match   32.5%; Score 583.5; DB 4; Length 212;
Best Local Similarity 68.3%; Pred. No. 6.5e-37;
Matches 129; Conservative 21; Mismatches 36; Indels 3; Gaps 3;

```
Qy     1 MACNNSGKDG-NASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATK 59
          ::|||||| | :|| |||||| |||||||||||||||||||||||:|||||||||||| |
Db    17 ISCNNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETLVLSIDELAKK 76

Qy    60 AIGKKI-GNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKN-EELKEKIDTAKQCSTEF 117
          ||||:|  |||| | :||:: |||||||||||:||| |||| |::|||:|:|| ||| ||
Db    77 AIGQKIDNNNGLAALNNQNGSLLAGAVAISTLITEKLSKLKNLEELKTEIAKAKKCSEEF 136

Qy   118 TNKLKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNA 177
          |||||:||| ::||  | ||||| |:||| |:|| |||||| ||:|:| ||:||||||||
Db   137 TNKLKSGHADLGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGLLKAAQVALTNS 196

Qy   178 VKELTSPIV 186
          ||||||:|:|
Db   197 VKELTSPVV 205
```

RESULT 13
US-09-196-293-11
Sequence 11, Application US/09196293
Patent No. 6183755
GENERAL INFORMATION:
APPLICANT: Fuchs, Renate
APPLICANT: Motz, Manfred
APPLICANT: Soutscheck, Erwin
APPLICANT: Wilske, Bettina
APPLICANT: Preac-Mursic, Vera
TITLE OF INVENTION: Active proteins from Borrelia
FILE REFERENCE: 738.001US2
CURRENT APPLICATION NUMBER: US/09/196,293
CURRENT FILING DATE: 1998-11-19
EARLIER APPLICATION NUMBER: US 08/209,603
EARLIER FILING DATE: 1994-03-10
EARLIER APPLICATION NUMBER: US 07/862,535
EARLIER FILING DATE: 1992-06-19
EARLIER APPLICATION NUMBER: WO PCT/EP90/02282
EARLIER FILING DATE: 1990-12-21
EARLIER APPLICATION NUMBER: DE P39 42 728.5
EARLIER FILING DATE: 1989-12-22
EARLIER APPLICATION NUMBER: DE P40 18 988.0
EARLIER FILING DATE: 1990-06-13
NUMBER OF SEQ ID NOS: 16
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 11
LENGTH: 212
TYPE: PRT
ORGANISM: Borrelia burgdorferi
US-09-196-293-11

Query Match   32.2%; Score 578.5; DB 4; Length 212;
Best Local Similarity 68.3%; Pred. No. 1.5e-36;
Matches 129; Conservative 19; Mismatches 38; Indels 3; Gaps 3;

```
Qy     1 MACNNSGKDG-NASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATK 59
          ::|||||| | :|| |||||| |||||||||||||||||||||||:|||||||||||| |
Db    17 ISCNNSGKVGILTSTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETLVLSIDELAKK 76

Qy    60 AIGKKI-GNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKN-EELKEKIDTAKQCSTEF 117
          ||||:|  |||| | :||:: |||||||||||:||| |||| |::|||:|:|| ||| ||
Db    77 AIGQKIDNNNGLAALNNQNGSLLAGAVAISTLITEKLSKLKNLEELKTEIAKAKKCSEEF 136

Qy   118 TNKLKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNA 177
          |||||:||| ::||  | ||||| |:||| |:|| |||||| ||:|:| ||:||||||||
Db   137 TNKLKSGHADLGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGLLKAAQVALTNS 196

Qy   178 VKELTSPIV 186
          ||||||:|:|
Db   197 VKELTSPVV 205
```

RESULT 14
US-08-209-603E-11
Sequence 11, Application US/08209603E
Patent No. 6248538
GENERAL INFORMATION:
APPLICANT: FUCHS, RENATE
APPLICANT: MOTZ, MANFRED
APPLICANT: SOUTSCHECK, ERWIN
APPLICANT: WILSKE, BETTINA
APPLICANT: PREAC-MURSIC, VERA
TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PROTEINS
TITLE OF INVENTION: FROM BORRELIA BURGDORFERI
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROOKS HAIDT HAFFNER & DELAHUNTY
STREET: 99 PARK AVENUE
CITY: NEW YORK
STATE: NY
COUNTRY: USA
ZIP: 10016
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" FLOPPY DISC
COMPUTER: AT&T - IBM COMPATIBLE
OPERATING SYSTEM: MS-DOS Version 6.2
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/209,603E
FILING DATE: 10-MAR-1994
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP90/02282
FILING DATE: 21-DEC-1990
APPLICATION NUMBER: US 07/862,535
FILING DATE: 19-JUN-1992
ATTORNEY/AGENT INFORMATION:
NAME: ROBINSON, WILLIAM R.
REGISTRATION NUMBER: 27,224
REFERENCE/DOCKET NUMBER: LKR-9217-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 697-3355
TELEFAX: (212) 557-5635
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 212
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
HYPOTHETICAL: N/A
ANTI-SENSE: N/A
FRAGMENT TYPE: N/A
ORIGINAL SOURCE:

```
;  ORGANISM: B. BURGDORFERI
;  IMMEDIATE SOURCE:
;    LIBRARY: DSM 5662
;  POSITION IN GENOME: N/A
;  FEATURE:
;  IDENTIFICATION METHOD: amino acid analysis
;  PUBLICATION INFORMATION:
;  ANTI-SENSE: N/A
;  FRAGMENT TYPE: INTERNAL
;  ORIGINAL SOURCE:
US-08-209-603E-11

Query Match         32.2%;  Score 578.5;  DB 4;  Length 212;
Best Local Similarity  68.3%;  Pred. No. 1.5e-36;
Matches  129;  Conservative  19;  Mismatches  38;  Indels   3;  Gaps   3;

Qy    1 MACNNSGKDG-NASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATK 59
          ::||||| |  ||||| ||||:|||||||||||||:||||||||||||:|||| ||||||
Db   17 ISCNNSGKVGLITSTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETLVLSIDELAKK 76

Qy   60 AIGKKI-GNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKN-EELKEKIDTAKQCSTEF 117
          |||:|| | |||  || :||| || || :|| |||| ||||| |  |||||||||| |||
Db   77 AIGQKIDNNNGLAALNNQNGSLLAGAYAISTLITEKLSKLKNLEELKTEIAKAKKCSEEF 136

Qy  118 TNKLKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNA 177
          |||| ||| | | || | | :  ||| || |||  ||||| |||||:|| ||| | | |
Db  137 TNKLKSGHADLGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGLLKAAQVALTNS 196

Qy  178 VKELTSPIV 186
          ||||:| |
Db  197 VKELTSPVV 205
```

```
;    TYPE: amino acid
;    TOPOLOGY: linear
;  MOLECULE TYPE: protein
US-08-235-836C-36

Query Match         31.3%;  Score 563;  DB 4;  Length 207;
Best Local Similarity  62.6%;  Pred. No. 2.2e-35;
Matches  117;  Conservative  34;  Mismatches  32;  Indels   4;  Gaps   3;

Qy    1 MACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKA 60
          ::|||||   | ||| ||||  |||||||||||||::||||||| ||||| |||| | |
Db   17 ISCNNSG--GDSASTNPDESAKGPNLTVISKKITDSNAFLLAVKEVEALLSSIDEL-SKA 73

Qy   61 IGKKIGNNG-LEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTN 119
          |||| |    |:|| |:|||:||||||||| |::||||:|  |||| |||||||| || |
Db   74 IGKKIKNDGTLDNEANRNESLLAGAYEISKLITQKLSVLNSEELKKIKEAKDCSEKFTT 133

Qy  120 KLKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVK 179
          ||| | | |||  | ::|| |||| || |||| ||||| ||||::|:: |||| | |
Db  134 KLKDSHAELGIQSVQDDNAKKAILKTHGTKDKGAKELEELFKSLESLSKAAQAALTNSVK 193

Qy  180 ELTSPIV 186
          ||:|:|
Db  194 ELTNPVV 200
```

Search completed: March 18, 2002, 09:55:35
Job time: 348 sec

RESULT 15
US-08-235-836C-36
Sequence 36, Application US/08235836C
Patent No. 6248562

GENERAL INFORMATION:
APPLICANT:
  APPLICANT: Dunn, John J.
  APPLICANT: Luft, Benjamin J.
TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising
TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor
NUMBER OF SEQUENCES: 144
CURRENT APPLICATION DATA:
COMPUTER READABLE FORM:
  MEDIUM TYPE: Floppy disk
  COMPUTER: IBM PC compatible
  OPERATING SYSTEM: PC-DOS/MS-DOS
  SOFTWARE: PatentIn Release #1.0, Version #1.25
CORRESPONDENCE ADDRESS:
  ADDRESSEE: Brookhaven National Laboratory
  STREET:
  CITY: Upton
  STATE: NY
  COUNTRY: USA
  ZIP: 11973
  APPLICATION NUMBER: US/08/235,836C
  FILING DATE: 29-APR-1994
  CLASSIFICATION: 435
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 08/148,191
  FILING DATE: 01-11-93
ATTORNEY/AGENT INFORMATION:
  NAME: Bogosian, Margaret C.
  REGISTRATION NUMBER: 25,324
  REFERENCE/DOCKET NUMBER: BNL93-28A
TELECOMMUNICATION INFORMATION:
  TELEPHONE: (516) 282-7338
  TELEFAX: (516) 282-3729
INFORMATION FOR SEQ ID NO: 36:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 207 amino acids
```

OM protein - protein search, using sw model

```
                    GenCore version 4.5
               Copyright (c) 1993 - 2000 Compugen Ltd.
```

Run on: March 18, 2002, 09:56:48 ; Search time 68.77 Seconds
                                            (without alignments)
                                            407.623 Million cell updates/sec

Title: 1798
Perfect score: US-09-596-746A-42
Sequence: 1 MACNNSGKDGNASANSADES...........AVENLAKAAKEMLANSVKEL 368

Scoring table:
                 BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

ol number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   PIR.68:*
                 1: pir1:*
                 2: pir2:*
                 3: pir3:*
                 4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	901	50.1	209	2	I40142	outer surface prot
2	876	48.7	193	2	S70287	outer surface prot
3	863	48.0	191	2	S69932	outer surface prot
4	826	45.9	178	2	I40277	outer surface prot
5	823	45.8	178	2	I40153	outer surface prot
6	739	41.1	178	2	I40125	outer surface prot
7	648.5	36.1	210	2	I40104	outer surface prot
8	646	35.9	210	2	I40124	outer surface prot
9	644.5	35.8	211	2	S69927	outer surface prot
10	642	35.7	210	2	I40145	outer surface prot
11	640	35.6	193	2	G70218	outer surface prot
12	639.5	35.6	193	2	S70279	outer surface prot
13	637	35.4	210	2	S70280	outer surface prot
14	635	35.3	211	2	I40144	outer surface prot
15	627	34.9	203	2	S69918	outer surface prot
16	625	34.8	211	2	I40108	outer surface prot
17	622	34.6	193	2	S70286	outer surface prot
18	616	34.3	209	2	S69917	outer surface prot
19	613	34.1	191	2	S70284	outer surface prot
20	612.5	34.1	194	2	S70277	outer surface prot
21	612.5	34.0	212	2	S69922	outer surface prot
22	602	33.5	194	2	S70268	outer surface prot
23	601.5	33.5	179	2	S54222	outer surface prot
24	600.5	33.4	212	2	I40143	outer surface prot
25	600.5	33.4	177	2	I40129	outer surface prot
26	594	33.0	209	2	S70254	outer surface prot
27	591	32.9	178	2	I40281	outer surface prot
28	589	32.8	193	2	I40106	outer surface prot
29	588.5	32.7	190	2	S70274	outer surface prot
30	588.5	32.7	212	2	S70261	outer surface prot
31	586	32.6	193	2	S20543	outer surface prot
32	586	32.6	211	2	S70276	outer surface prot
33	582	32.4	211	2	S69932	outer surface prot
34	579.5	32.2	212	2	I40277	outer surface prot
35	578.5	32.2	196	2	S69921	outer surface prot
36	578.5	32.2	212	2	S70290	outer surface prot
37	577	32.1	211	2	I40279	outer surface prot
38	575	32.0	193	2	S70265	outer surface prot
39	569	31.6	207	2	I40271	outer surface prot
40	567.5	31.6	192	2	S70285	outer surface prot
41	567	31.5	200	2	I40122	outer surface prot
42	567	31.5	207	2	S69919	outer surface prot
43	567	31.5	207	2	S37727	outer surface prot
44	565.5	31.5	194	2	S70289	outer surface prot
45	564	31.4	207	2	I40276	outer surface prot
```

ALIGNMENTS

RESULT   1
I40142
outer surface protein C precursor - Lyme disease spirochete
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 21-Jul-2000
C;Accession: I40142; S70283
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-209 <RES>
A;Cross-references: EMBL:U04240; NID:g2314877; PIDN:AAC45538.1; PID:g434658
A;Reference number: I40142; MUID:94314484
A;Accession: I40142
R;Stevenson, B.; Bockenstedt, L.K.; Barthold, S.W.
A;Title: Expression and gene sequence of outer surface protein C of Borrelia burgdorf
A;Reference number: I40142; MUID:94314484
R;Livey, I.; Gibbs, C.P.; Schuster, R.; Dorner, F.
A;Title: Evidence for lateral transfer and recombination in OspC variation in Lyme di
Infect. Immun. 62, 3568-3571, 1994
Mol. Microbiol. 18, 257-269, 1995
A;Reference number: S70255; MUID:96296448
A;Accession: S70283
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 19-209 <LIV>
A;Cross-references: EMBL:L42894; NID:g858722; PIDN:AAB37002.1; PID:g1695219
A;Experimental source: strain 28691
C;Genetics:
A;Gene: ospC
C;Superfamily: Lyme disease spirochete surface protein C

Query Match   50.1%;   Score 901;   DB 2;   Length 209;
Best Local Similarity   98.9%;   Pred. No. 4.4e-36;
Matches 184;   Conservative   2;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   1 MACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKA 60
            ::|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   17 ISCNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKA 76

Qy   61 IGKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNK 120
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   77 IGKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNK 136

Qy   121 LKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKE 180
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   137 LKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKE 196

Qy   181 LTSPIV 186
            ||||||
Db   197 LTSPIV 202

RESULT   2

```
S70287
outer surface protein C - Lyme disease spirochete
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 26-May-2000
C;Accession: S70287
R;Livey, I.; Gibbs, C.P.; Schuster, R.; Dorner, F.
Mol. Microbiol. 18, 257-269, 1995
A;Title: Evidence for lateral transfer and recombination in OspC variation in Lyme disea
A;Reference number: S70287; MUID:96296448
A;Accession: S70287
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-193 <LIV>
A;Cross-references: EMBL:L42895; NID:g858723; PIDN:AAB37003.1; PID:g1695220
A;Experimental source: strain 28354
A;Gene: ospC
C;Genetics:
C;Superfamily: Lyme disease spirochete surface protein C

Query Match       48.7%;  Score 876;  DB 2;  Length 193;
Best Local Similarity  100.0%;  Pred. No. 6e-35;
Matches  180;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy  189 NNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDELATKAIGKK 248
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2 NNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDELATKAIGK 61

Qy  249 KIQONGGLAVEAGHNGTLLAGAYTISKLITQKLIDGLKNSEKLKEKIENAKKCSEDFTKKL 308
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   62 KIQONGGLAVEAGHNGTLLAGAYTISKLITQKLIDGLKNSEKLKEKIENAKKCSEDFTKKL 121

Qy  309 EGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLAKAAKEMLANSVKEL 368
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  122 EGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLAKAAKEMLANSVKEL 181


RESULT   3
I40153
outer surface protein C - Lyme disease spirochete (fragment)
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 26-May-2000
C;Accession: I40153
R;Fung, B.P.; McHugh, G.L.; Leong, J.M.; Steere, A.C.
Infect. Immun. 62, 3213-3221, 1994
A;Title: Humoral immune response to outer surface protein C of Borrelia burgdorferi in
A;Reference number: I40153; MUID:94314437
A;Accession: I40153
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-191 <RES>
A;Cross-references: EMBL:U08284; NID:g469561; PIDN:AAA21460.1; PID:g469562
A;Genetics:
A;Gene: ospC
C;Superfamily: Lyme disease spirochete surface protein C

Query Match       48.0%;  Score 863;  DB 2;  Length 191;
Best Local Similarity   99.4%;  Pred. No. 2.4e-34;
Matches  178;  Conservative   0;  Mismatches   1;  Indels   0;  Gaps   0;

Qy  190 NSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDELATKAIGKK 249
        |||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 NSGKGGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDELATKAIGKK 60

Qy  250 IQONGGLAVEAGHNGTLLAGAYTISKLITQKLIDGLKNSEKLKEKIENAKKCSEDFTKKLE 309
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 IQONGGLAVEAGHNGTLLAGAYTISKLITQKLIDGLKNSEKLKEKIENAKKCSEDFTKKLE 120

Qy  310 GEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLAKAAKEMLANSVKEL 368
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 GEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLAKAAKEMLANSVKEL 179


RESULT   4
I40125
outer surface protein C - Lyme disease spirochete (strain MUL) (fragment)
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 26-May-2000
C;Accession: I40125; S54195
R;Theisen, M.; Borre, M.; Mathiesen, M.J.; Mikkelsen, B.; Lebech, A.M.; Hansen, K.
J. Bacteriol. 177, 3036-3044, 1995
A;Title: Evolution of the Borrelia burgdorferi outer surface protein OspC.
A;Reference number: I40125; MUID:95286481
A;Accession: I40125
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-178 <RES>
A;Cross-references: EMBL:X84779; NID:g793817; PIDN:CAA59250.1; PID:g793818
A;Gene: ospC
C;Superfamily: Lyme disease spirochete surface protein C

Query Match       45.9%;  Score 826;  DB 2;  Length 178;
Best Local Similarity   98.8%;  Pred. No. 1.2e-32;
Matches  171;  Conservative   0;  Mismatches   2;  Indels   2;  Gaps   0;

Qy  196 NTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDELATKAIGKKIQONGG 255
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 NTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDELATKAIGKKIQONGG 60

Qy  256 LAVEAGHNGTLLAGAYTISKLITQKLIDGLKNSEKLKEKIENAKKCSEDFTKKLEGEHAQL 315
        ||||||||||| |||||||||||||||||||||||| |||||||:||||||||||||||||
Db   61 LAVEAGHNGTSLLAGAYTISKLITQKLIDGLKNSEKLTEKIENAKRCSEDFTKKLEGEHAQL 120

Qy  316 GIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLAKAAKEMLANSVKEL 368
        |||||||||||||||||||||||||||||||||||||| |||||||||||||
Db  121 GIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLPKAAKEMLANSVKEL 173


RESULT   5
I40104
outer surface protein C - Lyme disease spirochete (strain 272) (fragment)
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 26-May-2000
C;Accession: I40104; S54184
R;Theisen, M.; Borre, M.; Mathiesen, M.J.; Mikkelsen, B.; Lebech, A.M.; Hansen, K.
J. Bacteriol. 177, 3036-3044, 1995
A;Title: Evolution of the Borrelia burgdorferi outer surface protein OspC.
A;Reference number: I40104; MUID:95286481
A;Accession: I40104
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-178 <RES>
A;Cross-references: EMBL:X84785; NID:g793787; PIDN:CAA59256.1; PID:g793788
A;Gene: ospC
C;Superfamily: Lyme disease spirochete surface protein C

Query Match       45.8%;  Score 823;  DB 2;  Length 178;
Best Local Similarity   98.3%;  Pred. No. 1.7e-32;
Matches  170;  Conservative   0;  Mismatches   1;  Indels   2;  Gaps   0;

Qy  196 NTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDELATKAIGKKIQONGG 255
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 NTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDELATKAIGKKIQONGG 60

Qy  256 LAVEAGHNGTLLAGAYTISKLITQKLIDGLKNSEKLKEKIENAKKCSEDFTKKLEGEHAQL 315
        ||||||||||||||||||||||||||||||||||| |||||||:||||||||||||||||
Db   61 LAVEAGHNGTLLAGAYTISKLITQKLIDGLKNSEKLTEKIENAKRCSEDFTKKLEGEHAQL 120

Qy  316 GIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLAKAAKEMLANSVKEL 368
        |||||||||||||||||||||||||||||||||||||| |||||||||||||
Db  121 GIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLPKAAKEMLANSVKEL 173
```

RESULT 6

I40124
outer surface protein C - Lyme disease spirochete (strain KIPP) (fragment)
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 26-May-2000
C;Accession: I40124; S54194
A;Molecule type: DNA
A;Residues: 1-178 <RES>
R;Thelsen, M.; Borre, M.; Mathiesen, M.J.; Mikkelsen, B.; Lebech, A.M.; Hansen, K.
J. Bacteriol. 177, 3036-3044, 1995
A;Title: Evolution of the Borrelia burgdorferi outer surface protein OspC.
A;Reference number: I40124
A;Accession: I40124
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-178 <RES>
A;Cross-references: EMBL:X84782; NID:g793815; PIDN:CAA59253.1; PID:g793816
C;Superfamily: Lyme disease spirochete surface protein C

Query Match   41.1%;   Score 739;   DB 2;   Length 178;
Best Local Similarity   89.0%;   Pred. No. 1.5e-28;
tches 154;   Conservative   6;   Mismatches 13;   Indels   0;   Gaps   0;

Qy   196 NTSANSADESVKGPNLTEISKKITTESNAVVLAVKEIETLLASIDELATKAIGKKIQQNGG 255
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 NTSANSADESVKGPNLTEISKKITTESNAVVLAVKEIETLLASIDELATKAIGKKIQQNGG 60

Qy   256 LAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKEKIENAKKCSEDFTKKLEGEHAQL 315
         |||||||||||||||||||||||||||||:.|||||||||||||||||||||||||||||
Db   61 LAVEAGHNGTSLAGAYTISKLITQKIRWIENSEKITGKIENAKKCSEDFTNKLEGEHAQL 120

Qy   316 GIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLAKAAKEMLANSVKEL 368
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GTENVTDENADKAILTDAAKDKGAAELAKLFKAVENLPQAAKEMLANSVKEL 173

RESULT 7

S69927
outer surface protein C precursor - Lyme disease spirochete (strain Pka)
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
A;Variety: strain Pka
C;Date: 15-Feb-1997 #sequence_revision 27-Feb-1997 #text_change 26-May-2000
C;Accession: S69927; S72669
R;Jauris-Helpke, S.; Liegl, G.; Preac-Mursic, V.; Roessler, D.; Schwab, E.; Soutschek, E.
J. Clin. Microbiol. 33, 1860-1866, 1995
A;Title: Molecular analysis of genes encoding outer surface protein C (OspC) of Borrelia
A;Reference number: I40047; MUID:95395018
A;Accession: S69927
A;Accession: S69927
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
esidues: 1-210 <JAU>
ross-references: EMBL:X69589
A;Experimental source: strain Pka
R;Jauris, S.
submitted to the EMBL Data Library, February 1994
A;Accession: S72669
A;Molecule type: DNA
A;Residues: 1-124, 'D',126-139,'E',141-210 <JAW>
A;Cross-references: EMBL:X69589
A;Genetics:
A;Gene: ospC
C;Superfamily: Lyme disease spirochete surface protein C

Query Match   36.1%;   Score 648.5;   DB 2;   Length 210;
Best Local Similarity   76.1%;   Pred. No. 3.2e-24;
Matches 137;   Conservative 14;   Mismatches 28;   Indels   1;   Gaps   1;

Qy   189 NNSGKGNTSANSADESVKGPNLTEISKKITTESNAVVLAVKEIETLLASIDELATKAIGK 248
         ||||||||||||||||||||||||||||||||||||||||||:.||||||||||||||||
Db   20 NNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGK 79

Qy   249 KIQQNGGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKEKIENAKKCSEDFTKKL 308

Db   80 KIHQNNGLDTENNHNGSLLAGAYAISTLIKQKIDGLKN-EGLKEKIEAAKKCSETFTNKL 138

Qy   309 EGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLAKAAKEMLANSVKEL 368
         |||:.||||||| ||||||||||:.|||||||||||||||||||||:.|||||||||||||
Db   139 KDKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKEL 198

RESULT 8

I40145
outer surface protein C precursor - Lyme disease spirochete
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 21-Jul-2000
A;Molecule type: DNA
A;Residues: 1-211 <RES>
A;Cross-references: EMBL:U04282; NID:g2314881; PIDN:AAC45540.1; PID:g434666
R;Stevenson, B.; Barthold, S.W.
FEMS Microbiol. Lett. 124, 367-372, 1994
A;Title: Expression and sequence of outer surface protein C among North American isol
A;Reference number: I40143; MUID:95154673
A;Accession: I40145
A;Accession: I40145
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-211 <RES>
A;Cross-references: EMBL:U04282; NID:g2314881; PIDN:AAC45540.1; PID:g434666
A;Gene: ospC
C;Genetics:
C;Superfamily: Lyme disease spirochete surface protein C

Query Match   35.9%;   Score 646;   DB 2;   Length 211;
Best Local Similarity   72.3%;   Pred. No. 4.2e-24;
Matches 136;   Conservative 18;   Mismatches 32;   Indels   2;   Gaps   2;

Qy   1 MACNNSGKDGN-ASANSADESVKGPNLTEISKKITTESNAVVLAVKEIETLLASIDELATK 59
         :.||||||||| |:||||||||||||||||||||||:.|||||||||:.|||||||||||
Db   17 ISCNNSGKDGNAASTNPADESVKGPNLTEISKKITDSNAVLLAVKEVEGALLTSIDELATK 76

Qy   60 AIGKKI-GNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFT 118
         |||||| ||:.|||:.|:|||:.||||||:.||||:.|||:.||||||:.||||||||||
Db   77 AIGKKIHQNNGLDTENNHNGSLLAGAYAISTLITQKLGGLKNEELKEKTAAVKKCSEEFT 136

Qy   119 NKLKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAV 178
         |||||||:.|||||||||||:.|||||||:.|||||||||||||||||::.|||:.||||
Db   137 NKLKSSHTELGKQDAQDDDAKKAILRTHNTKDKGAEELDKLFKAVENLSKSKAAKEMLSNSV 196

Qy   179 KELTSPIV 186
         |||::.|:.|
Db   197 KELTSPVV 204

RESULT 9

G70218
outer surface protein C - Lyme disease spirochete
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 21-Jul-2000
C;Accession: G70218; I40269; S37726; S70281
R;Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; Wh
son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vu
; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997
A;Authors: Smith, H.O.; Venter, J.C.
A;Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.
A;Reference number: A70100; MUID:98065943
A;Accession: G70218
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-210 <KLE>
A;Cross-references: GB:AE000792; NID:g3253098; PIDN:AAC66329.1; PID:g2689901; TIGR:BB
A;Experimental source: strain B31
R;Fukunaga, M.; Hamase, A.
J. Clin. Microbiol. 33, 2415-2420, 1995
A;Title: Outer surface protein C gene sequence analysis of Borrelia burgdorferi sensu
A;Reference number: I40269; MUID:96025162
A;Accession: I40269
A;Accession: I40269

A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-210 <RES>
A:Cross-references: GB:D49497; NID:g7070092, PIDN:BAA08457.1; PID:g769684
R:Jauris-Heipke, S.; Fuchs, R.; Motz, M.; Preac-Mursic, V.; Schwab, E.; Soutschek, E.; W
Med. Microbiol. Immunol. 182, 37-50, 1993
A:Title: Genetic heterogeneity of the genes coding for the outer surface protein C (OspC)
A:Reference number: S37726; MUID:93268136
A:Accession: S37726
A:Status: preliminary
A:Molecule type: DNA
R:Livey, I.; Gibbs, C.P.; Schuster, R.; Dorner, F.
A:Cross-references: <JAU>
A:Residues: 1-210
Mol. Microbiol. 18, 257-269, 1995
A:Title: Evidence for lateral transfer and recombination in OspC variation in Lyme disea
A:Reference number: S70255; MUID:96296448
Accession: S70281
Atus: nucleic acid sequence not shown
olecule type: DNA
A:Residues: 19-210 <LIV>
A:Cross-references: EMBL:L42887; NID:g858715; PIDN:AAB36995.1; PID:g1695212
A:Experimental source: strain lp2
C:Gene: ospC
C:Superfamily: Lyme disease spirochete surface protein C

Query Match   35.8%;   Score 644.5; DB 2;   Length 210;
Best Local Similarity   75.6%;   Pred. No. 4.9e-24;
Matches 136; Conservative 15; Mismatches 28;   Indels 1;   Gaps 1;

Qy 189 NNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDELATKAIGK 248
        ||||||||||||||||||||||||||||||||||| |||:|||| |||| |||||||||
Db   20 NNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEVEALLSSIDEIAAKAIGK 79

Qy 249 KIQQNGLAVEAGHNGTLLAGAYTISKLITQKLDGILKNSEKLKEKIENAKKCSEDFTKKL 308
        |||||||||||||||||||||||||||||||||||| |||| ||||||||||||| ||
Db   80 KIHQNNGLDTENNHNGSLLAGAYAISTLIKOKLDGLKN-EGLKEKIDAARKCSEFFTNKL 138

Qy 309 EGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLAKAAKEMLANSVKEL 368
        :|| ||| || ||||||||||||||||:| |||||||||| |||||||||||||||||
Db 139 KEKHTDLGKEGVTDADAKEALLKTNGTKTGAEELGKLFESVEVLSKAAKEMLANSVKEL 198

RESULT 10
S70279
surface protein C - Lyme disease spirochete
Species: Borrelia burgdorferi (Lyme disease spirochete)
C:Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 26-May-2000
C:Accession: S70279
A:Genetics:
A:Gene: ospC
C:Superfamily: Lyme disease spirochete surface protein C

Query Match   35.7%;   Score 642;   DB 2;   Length 193;
Best Local Similarity   72.6%;   Pred. No. 5.8e-24;
Matches 135; Conservative 16; Mismatches 33;   Indels 2;   Gaps 2;

Qy   3 CNNSGKDGN-ASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKAI 61
        | ||||||| |:|||||||||||||||||||||||||||||| |||| |||||||||||
Db   1 CNNSGKDGNAASTNPADESVKGPNLTEISKKITDSNIVVLAVKEVEALLTSIDELATKAI 60

RESULT 11
S70280
outer surface protein C - Lyme disease spirochete
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C:Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 26-May-2000
C:Accession: S70280
R:Livey, I.; Gibbs, C.P.; Schuster, R.; Dorner, F.
Mol. Microbiol. 18, 257-269, 1995
A:Title: Evidence for lateral transfer and recombination in OspC variation in Lyme di
A:Reference number: S70255; MUID:96296448
A:Accession: S70280
A:Residues: 1-193 <LIV>
A:Cross-references: EMBL:L42868; NID:g858735; PIDN:AAB37011.1; PID:g1695226
A:Experimental source: strain ZS7
C:Genetics:
A:Gene: ospC
C:Superfamily: Lyme disease spirochete surface protein C

Query Match   35.6%;   Score 640;   DB 2;   Length 193;
Best Local Similarity   76.2%;   Pred. No. 7.2e-24;
Matches 138; Conservative 16; Mismatches 25;   Indels 2;   Gaps 2;

Qy 189 NNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDELATKAIGK 248
        ||||||||||||||||||||||||||||||||||| ||||||| ||||:|||||||||
Db   2 NNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEVEALLSSIDELA-KAIGK 60

Qy 249 KIQQNGLAVEAGHNGTLLAGAYTISKLITQKLDGILKNSEKLKEKIENAKKCSEDFTKKL 308
        |||||||||| || || |||||||||||||||||||||||||||||||||||||| || ||
Db   61 KIKNDGSLGDEANHNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEFFSTKL 120

Qy 309 EGEHAQLGIENVTDENAKKAILITDAA-KDKGAAELEKLFKAVENLAKAAKEMLANSVKE 367
        ||||||||||||||||||| |||||| |||||||||||||||:| ||||||||||||||
Db 121 KDNIHAQLGIGGVTDENAKKAILANNAAGKDKGVEELEKLSGSLESLSKAAKEMLANSVKE 180

Qy 368 L 368
        |
Db 181 L 181

RESULT 12
I40144
outer surface protein C precursor - Lyme disease spirochete
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C:Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 21-Jul-2000
C:Accession: I40144; S70282
R:Stevenson, B.; Barthold, S.W.
FEMS Microbiol. Lett. 124, 367-372, 1994
A:Title: Expression and sequence of outer surface protein C among North American isol
A:Reference number: I40143; MUID:95154673
A:Accession: I40144
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-210 <RES>
A:Cross-references: EMBL:U04281; NID:g434663; PIDN:AAC43297.1; PID:g434664
R:Livey, I.; Gibbs, C.P.; Schuster, R.; Dorner, F.
Mol. Microbiol. 18, 257-269, 1995

RESULT (continued) — left column:

RESULT 9
S70279 / surface protein C

A:Accession: X69596; NID:g311391; PIDN:CAA49306.1; PID:g311392

Qy 62 GKKI-GNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNK 120
       ||||| : |||||||||:||: : |:||| ||||||||||||||||||||| ||||||||
Db 61 GKKIHQNNGLDTENNHNGSLLAGAYAISTLITOKLGGGIKNEELKEKIAAAVKCSEEFTNK 120

Qy 121 LKSEHAVLGLDNLTDDNAQRAILKHHANKDKGAAELEKLFKAVENLSKKAAQDTLKNAVKE 180
       ||| ::||||| ||||||:|:|||||||||||||||||||||| |||:| |:||||||
Db 121 LKSSHTELGKQDKAQDDDKKAILRTHNTKDKGAEELDKLFKPVENLSKAAKEMLSNSVKE 180

Qy 181 LTSPIV 186
       |||||:
Db 181 LTSPVV 186

Query Match   35.8%;   Score 642;   DB 2;
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-193 <LIV>
A:Accession: S70279
A:Reference number: S70255; MUID:9296648
A:Title: Evidence for lateral transfer and recombination in OspC variation in Lyme disea
Mol. Microbiol. 18, 257-269, 1995
A:Cross-references: EMBL:L42898; NID:g858729; PIDN:AAB37007.1; PID:g1695223
C:Date: 12-Feb-1998
C:Species: Borrelia burgdorferi (Lyme disease spirochete)

A:Title: Evidence for lateral transfer and recombination in OspC variation in Lyme disea
A:Reference number: S70282
A:Accession: S70282
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 19-210 <LIV>
A:Cross-references: EMBL:L42893; NID:g858721; PIDN:AAB37001.1; PID:g1695218
A:Experimental source: strain 297
C:Genetics:
A:Gene: ospC
C:Superfamily: Lyme disease spirochete surface protein C

Query Match   35.6%; Score 639.5; DB 2; Length 210;
Best Local Similarity   74.4%; Pred. No. 8.4e-24;
Matches 134; Conservative 20; Mismatches 25; Indels   1; Gaps   1;

Qy 189 NNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDELATKAIGK 248
        |||||||||||||||||||||||||||||||||||||:|||||||:||||||||||||||
Db   20 NNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLTSIDELA-KAIGK 78

Qy 249 KIQQNGLAVEAGHNGTLLAGAYTISKLITQKLIDGLKNSEKLKEKIENAKKCSEDFTKKL 308
        |: || |||| || |||||||||||||:|||:||| ||||:|||||:|| ||||||||
Db   79 KIKNDVSLDNEADHNGSLISGAYLISTLITQKLISAIKDSGELKAEIEKRAKKCSEEFTKL 138

Qy 309 EGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLAKAAKEMLANSVKEL 368
        |||| | ||||:|||||||||||: | |||||:|||||||||||||:|| |||||||||
Db 139 KGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNLSKAAKEMLTNSVKEL 198

Db 139 KDNHAQLGIQGVTDDNAKKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANSVKE 198

Qy 368 L 368
        |
Db 199 L 199

RESULT 13
S69918
outer surface protein C precursor - Lyme disease spirochete (strain PBre)
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
A;Variety: strain PBre
A;Accession: S69918; S72674; I40103
C;Date: 06-Dec-1996 #sequence_revision 14-Feb-1997 #text_change 26-May-2000
R;Jauris-Heipke, S.; Liegl, G.; Preac-Mursic, V.; Roessler, D.; Schwab, E.; Soutschek, E
J. Clin. Microbiol. 33, 1860-1866, 1995
A;Title: Molecular analysis of genes encoding outer surface protein C (OspC) of Borrelia
A;Reference number: 140047; MUID:95395018
A;Accession: S69918
A;Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-211 <JAV>
A:Cross-references: EMBL:X81522; NID:g872021; PIDN:CAA57242.1; PID:g872022
A:Experimental source: strain PBre
R:Roessler, D.
submitted to the EMBL Data Library, September 1994
reference number: S72674
accession: S72674
A;Molecule type: DNA
A;Residues: 1-152,'E',154-211 <ROE>
A;Cross-references: EMBL:X81522; NID:g872021; PIDN:CAA57242.1; PID:g872022
C;Genetics:
A;Gene: ospC
C;Superfamily: Lyme disease spirochete surface protein C

Query Match   35.4%; Score 637; DB 2; Length 211;
Best Local Similarity   75.8%; Pred. No. 1.1e-23;
Matches 137; Conservative 17; Mismatches 25; Indels   2; Gaps   2;

Qy 189 NNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDELATKAIGK 248
        |||||||||||||||||||||||||||||||||||||:|||||||:||||||||||||||
Db   20 NNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVEALLSSIDELA-KAIGK 78

Qy 249 KIQQNGLAVEAGHNGTLLAGAYTISKLITQKLIDGLKNSEKLKEKIENAKKCSEDFTKKL 308
        |: || |||| ||:|||||||||:|||||||:|| |||||:| |||:| || |||||
Db   79 KIKNDGSLGDEANHNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEEFSTKL 138

Qy 309 EGEHAQLGIENVTDENAKKAILITDAA-KDKGAAELEKLFKAVENLAKAAKEMLANSVKE 367
        :||||||||||:||||||| ||| | | |||| ||||||||||:|||:|||||||||||||
Db 139 KIKNDGSLGDEANHNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEEFSTKL 138

Qy 309 EGEHAQLGIENVTDENAKKAILITDAA-KDKGAAELEKLFKAVENLAKAAKEMLANSVKE 367
        :|||||||||||||||||| ||| | | |||| |||||||||||:|||:||||||||||
Db 139 KDNHAQLGIQGVTDDNAKKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANSVKS 198

Qy 368 L 368
        |
Db 199 L 199

RESULT 14
I40108
outer surface protein C - Lyme disease spirochete (strain DK7) (fragment)
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
J. Clin. Microbiol. 31, 2570-2576, 1993
R;Theisen, M.; Frederiksen, B.; Lebech, A.M.; Vuust, J.; Hansen, K.
A;Title: Polymorphism in ospC gene of Borrelia burgdorferi and immunoreactivity of Os
A;Reference number: I40108; S34176
C;Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 26-May-2000
A;Accession: I40108
A;Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-203 <RES>
A:Cross-references: EMBL:X73625; NID:g313275; PIDN:CAA52004.1; PID:g313276
A;Accession: I40105; MUID:94075528
C;Genetics:
A;Gene: ospC
C;Superfamily: Lyme disease spirochete surface protein C

Query Match   35.3%; Score 635; DB 2; Length 203;
Best Local Similarity   75.8%; Pred. No. 1.3e-23;
Matches 137; Conservative 16; Mismatches 26; Indels   2; Gaps   2;

Qy 189 NNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDELATKAIGK 248
        |||||||||||||||||||||||||||||||||||||:|||||||:||||||||||||||
Db   20 NNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVEALLSSIDELA-KAIGK 78

Qy 249 KIQQNGLAVEAGHNGTLLAGAYTISKLITQKLIDGLKNSEKLKEKIENAKKCSEDFTKKL 308
        |: || |||| ||:|||||||||:|||||||:|| |||||:| |||:| || |||||
Db   79 KIKNDGSLGDEANHNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEEFSTKL 138

Qy 309 EGEHAQLGIENVTDENAKKAILITDAA-KDKGAAELEKLFKAVENLAKAAKEMLANSVKE 367
        :||||||||||:||||||| ||| | | |||| |||||||||||:|||:||||||||||
Db 139 KDNHAQLGIQGVTDENAKKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANSVKS 198

Qy 368 L 368
        |
Db 199 L 199

RESULT 15
I40268
outer surface protein C precursor - Lyme disease spirochete
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 26-May-2000
A;Accession: I40268
R;Margolis, N.; Hogan, D.; Cieplak, W.
Gene 143, 105-110, 1994
A;Title: Homology between Borrelia burgdorferi OspC and members of the family of Borr
A;Reference number: I40268; MUID:94259285
A;Accession: I40268
A;Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-211 <RES>
A:Cross-references: GB:L25413; NID:g495735; PIDN:AAA22956.1; PID:g495736
C;Genetics:
A;Gene: ospC
C;Superfamily: Lyme disease spirochete surface protein C

Query Match   34.9%; Score 627; DB 2; Length 211;
Best Local Similarity   72.8%; Pred. No. 3.3e-23;
Matches 131; Conservative 16; Mismatches 33; Indels   0; Gaps   0;

Qy 189 NNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDELATKAIGK 248

```
Db   20  NNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEVLLSSIDELAKKAIGK  79
             |||||||||||||||||||||||||||||||||||||||:|||:||||||||||||||:||  ||||
Qy  249  KIQQNGGLAVEAGHNGTLLAGAYTISKLIITQKLDGLKNSEKLKEKIENAKKCSEDFTKKL  308
             ||:|:| ||||:|  |||:|| :|:| ||||||||||||
Db   80  KIDQNNALGTLDNHNGSLLAGAYAISALITEKLSSSIKDSGELKAEIEKAKKCSESFTKKL  139

Qy  309  EGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLAKAAKEMLANSVKEL  368
             |:|||||||||:|||||||||||:| ||||:|||||||  ||||  ::|  |||||||||||
Db  140  SDNQAELGIENATIDDNAKKAILKTHNAKDKGAEELVKLSESVAGLLKAAQAILANSVKEL  199
```

us-09-596-746a-42.rapn

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2002, 09:58:28 ; Search time 81.02 seconds
    (without alignments)
    290.158 Million cell updates/sec

Title: US-09-596-746A-42
Perfect score: 1798
Sequence: 1 MACNNSGKDGNASANSADES..........AVENLAKAAKEMLANSVKEL 368

Scoring table: BLOSUM62
  Gapop 10.0 , Gapext 0.5
Searched: 303921 seqs, 63882009 residues

Total number of hits satisfying chosen parameters: 303921

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
  Maximum Match 100%
  Listing first 45 summaries

Database : Pending_Patents_AA_New:*
 1: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pep:*
 2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep:*
 3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep:*
 4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep:*
 5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep:*
 6: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep1:*
 7: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep:*
 8: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	% Query Match Length	DB ID	DB ID	Description
1	640	35.6	211	6	US-09-974-992-5	Sequence 5, Appli
2	588.5	32.7	212	6	US-09-974-992-7	Sequence 7, Appli
3	567	31.5	207	6	US-09-974-992-3	Sequence 3, Appli
4	176	9.8	837	7	US-09-815-242-5883	Sequence 5883, Appli
5	176	9.8	837	7	US-10-072-851-5883	Sequence 5883, Ap
6	176	9.8	875	7	US-09-815-242-13080	Sequence 13080, A
7	176	9.8	875	7	US-10-072-851-13080	Sequence 13080, A
8	170	9.5	1304	6	US-09-708-427-15045	Sequence 15045, A
9	170	9.5	1304	6	US-09-708-427-15044	Sequence 15044, A
10	169.5	9.4	1313	6	US-09-708-427-15044	Sequence 15044, A
11	169.5	9.4	2434	6	US-10-072-851-5835	Sequence 5835, A
12	169.5	9.4	2434	6	US-09-815-242-5835	Sequence 5835, A
13	169.5	9.4	6281	7	US-10-072-851-12996	Sequence 12996, A
14	168.5	9.4	1765	7	US-10-037-182-8	Sequence 8, Appli
15	168.5	9.4	1786	5	US-09-938-275-6	Sequence 6, Appli
16	168.5	9.4	1786	6	US-09-873-676-113	Sequence 113, App
17	168.5	9.4	1786	6	US-09-561-709B-9	Sequence 9, Appli
18	168.5	9.4	1786	7	US-10-037-182-6	Sequence 6, Appli
19	167.5	9.3	1013	6	US-09-614-150-9159	Sequence 9159, Ap
20	169.5	9.1	1881	6	US-09-708-427-19882	Sequence 7646, Ap
21	163.5	9.1	1881	7	US-09-708-427-19882	Sequence 15590, A
22	162	9.0	1014	6	US-09-708-427-19882	Sequence 15590, A
23	162	9.0	1018	6	US-09-708-427-19883	Sequence 19883, A
24	162	9.0	1269	6	US-09-708-427-19881	Sequence 19882, A
25	161.5	9.0	730	6	US-09-708-427-3731	Sequence 3731, Ap
26	161.5	9.0	805	6	US-09-708-427-3730	Sequence 3730, Ap
27	161.5	9.0	841	6	US-09-708-427-3729	Sequence 3729, Ap
28	158.5	8.8	5795	6	US-09-815-242-12610	Sequence 12610, A
29	158.5	8.8	5795	7	US-10-072-851-12610	Sequence 12610, A
30	155.5	8.6	2056	7	US-09-614-150-4824	Sequence 4824, Ap
31	155	8.6	1038	6	US-09-708-427-32624	Sequence 32624, A
32	155	8.6	1042	6	US-09-708-427-32623	Sequence 32623, A
33	155	8.6	1043	6	US-09-708-427-32622	Sequence 32622, A
34	155	8.6	2076	6	US-09-815-242-5815	Sequence 5815, Ap
35	155	8.6	2076	7	US-10-072-851-5815	Sequence 5815, Ap
36	155	8.6	2186	7	US-09-815-242-12913	Sequence 12913, A
37	155	8.6	2186	7	US-10-072-851-12913	Sequence 12913, A
38	155	8.6	2437	6	US-09-815-242-5834	Sequence 5834, Ap
39	155	8.6	2437	7	US-10-072-851-5834	Sequence 5834, Ap
40	154.5	8.6	1144	6	US-09-708-427-15046	Sequence 15046, A
41	153	8.5	1725	7	US-10-037-182-12	Sequence 12, Appl
42	153	8.5	1786	5	US-09-938-275-7	Sequence 7, Appl
43	153	8.5	1786	6	US-10-037-182-10	Sequence 10, Appl
44	152.5	8.5	630	6	US-09-742-096-5	Sequence 5, Appli
45	152	8.5	1128	7	US-10-032-585-7627	Sequence 7627, Ap

ALIGNMENTS

RESULT 1
US-09-974-992-5
Sequence 5, Application US/09974992
; GENERAL INFORMATION:
; APPLICANT: Mathiesen, Marianne J.
; APPLICANT: Theisen, Michael
; APPLICANT: Holm, Arne
; APPLICANT: Ostergaard, Soren
; TITLE OF INVENTION: Novel OspC-derived peptide fragments
; FILE REFERENCE: 459-666P
; CURRENT APPLICATION NUMBER: US/09/974,992
; CURRENT FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US/09/180,089
; PRIOR FILING DATE: 1999-05-13
; PRIOR APPLICATION NUMBER: PCT/DK97/00203
; PRIOR FILING DATE: 1997-05-02
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: Patentin Ver. 2.1
;
; SEQ ID NO 5
; LENGTH: 211
; TYPE: PRT
; ORGANISM: Borrelia burgdorferi
US-09-974-992-5

Query Match 35.6%; Score 640; DB 6; Length 211;
Best Local Similarity 76.2%; Pred. No. 8.2e-37;
Matches 138; Conservative 16; Mismatches 25; Indels 2; Gaps 2;

Qy 189 NNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDELATKAIGK 248
    |||||||||||||||:||||||||||||| |||| :| |||||||| | ||||||||
Db 20 NNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDELA-KAIGK 78

Qy 249 KIQONGGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKEKIENAKKCSEDFTKKL 308
    || : :| || | ||| :|||||||| |||||||| | | ||||| ||||||| ||
Db 79 KIKNDGSLGDDEANHNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAARKCSEEFSTKL 138

Qy 309 EGEHAQLGIENVTDENAKKAILITDAA-KDKGAAELEKLFKRAVENLAKAAKEMLANSVKE 367
     ||||||| :||||||||||||| :|| |||| | |||| ::|||||||||||||||||||
Db 139 KDNHAQLGIQGVTDENAKKAILKANAAGKDKGVEELEKLSGSLESLSKAAKEMLANSVKE 198

Qy 368 L 368
   |
Db 199 L 199

RESULT 2
US-09-974-992-7

; Sequence 7, Application US/09974992
; GENERAL INFORMATION:
; APPLICANT: Mathiesen, Marianne J.
; APPLICANT: Theisen, Michael
; APPLICANT: Holm, Arne
; APPLICANT: Ostergaard, Soren
; TITLE OF INVENTION: Novel OspC-derived peptide fragments
; FILE REFERENCE: 459-666P
; CURRENT APPLICATION NUMBER: US/09/974,992
; CURRENT FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: 09/180,089
; PRIOR FILING DATE: 1999-05-13
; PRIOR APPLICATION NUMBER: PCT/DK97/00203
; PRIOR FILING DATE: 1997-05-02
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 7
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Borrelia afzelii
US-09-974-992-7

Query Match   32.7%; Score 588.5; DB 6; Length 212;
Best Local Similarity 68.8%; Pred. No. 2.7e-33;
Matches 130; Conservative 20; Mismatches 36; Indels 3; Gaps 3;

```
Qy    1 MACNNSGKDG-NASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATK  59
          ::||||||| | ::|| |||| |||||||||||||| ||||| |||||||:||||  |
Db   17 ISCNNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVEVALLSSIDEL-SKA  76

Qy   60 AIGKKI-GNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKN-EELKEKIDTAKQCSTEF 117
         ||||| | |||||||| ||| ||| |||||||:||| |:||||  |
Db   77 AIGQKIDNNNGLAALNNQNGSLLAGAYAISTLITEKLSKLKNLEELKTEIAKAKKCSEEF 136

Qy  118 TNKLKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNA 177
         |||||| ||| || |||||||||||||||||||||||||:||  |
Db  137 TNKLKSGHADLGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGLLKAAQVALTNS 196

Qy  178 VKELTSPIV 186
         ||||||:|:|
Db  197 VKELTSPVV 205
```

RESULT 3
US-09-974-992-3
; Sequence 3, Application US/09974992
; GENERAL INFORMATION:
; APPLICANT: Mathiesen, Marianne J.
; APPLICANT: Theisen, Michael
; APPLICANT: Holm, Arne
; APPLICANT: Ostergaard, Soren
; TITLE OF INVENTION: Novel OspC-derived peptide fragments
; FILE REFERENCE: 459-666P
; CURRENT APPLICATION NUMBER: US/09/974,992
; CURRENT FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: 09/180,089
; PRIOR FILING DATE: 1999-05-13
; PRIOR APPLICATION NUMBER: PCT/DK97/00203
; PRIOR FILING DATE: 1997-05-02
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 3
; LENGTH: 207
; TYPE: PRT
; ORGANISM: Borrelia garinii
US-09-974-992-3

Query Match   31.5%; Score 567; DB 6; Length 207;
Best Local Similarity 63.1%; Pred. No. 7.8e-32;
Matches 118; Conservative 33; Mismatches 32; Indels 4; Gaps 3;

```
Qy    1 MACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKA  60
         :||||| |::::| |||| :: ||||||||||||||||||||||:| ||||||| :|| :||
Db   17 ISCNNSG--GDSASTNPDESAKGPNLTVISKKITDSNAPLLAVKEVEVALLSSIDEL-SKA  73

Qy   61 IGKKIGNNG-LEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTN 119
         ||||| :|:|| :|:||:|| :|||| | :|| |||:|| :|||||||:|| ||:| :||
Db   74 IGKKIKNDGTLDNEANRNESLIAGAYEISKLITQKLSVLNSEELKEKIKEAKDCSEKFTT 133

Qy  120 KIKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVK 179
         |:|:||| ||: :: |:|:|:|||| ||:|||||:||||||:::: |:|||||:| ::|:|
Db  134 KLKDSHAELGIQSVQDDNAKKAILKTHGTKDKGAKELEELFKSLESLSKAQAALTNSVK 193

Qy  180 ELTSPIV 186
         ||| |:|
Db  194 ELTNPVV 200
```

RESULT 4
US-09-815-242-5883
; Sequence 5883, Application US/09815242
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl I.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5883
; LENGTH: 837
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-5883

Query Match   9.8%; Score 176; DB 6; Length 837;
Best Local Similarity 23.1%; Pred. No. 0.00026;
Matches   93; Conservative 77; Mismatches 154; Indels 78; Gaps 20;

```
Qy    4 NNSGKDGNASANSADE---SVKGPNLTEISKKITESNAV--VLAVKE-VETLLASI---  53
         ||| | |||||:||:|   | |||:|| : | | :||:|  :| |||  |||||:||
Db  100 NHNLEVAKQNANTAIDGLTSLNGPQAKLKEQVGQATTLPNVQTVRDNAQTLNTAMKGLR 159

Qy   54 DELATKAIGKKIGNNGLEANQSKNTSLLSGAYAISDLIAE----KLNVLKNEELKEKIDT 109
         | :| | | ||| :| :| : |  | :|:|| |:  :| |    ||||:|| |||:|:||
Db  160 DSIANEAT-IKAGQNYTDASQNKQTDYNSAVTAAKAIIGQTTSPSMNAQEINQAKDQV-T 217

Qy  110 AKQCSTEFTNKLK-----SEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVE 164
         |||  ||: || |     |:| : :| |||:||||: :| | || : || | ||||| |:
Db  218 AKQQALNGQENLRTAQTNAKQHLNGLSDLTD--AQKDAVKRQI---EGATHVNEVTQAQN 272
```

Left column

```
Qy    165  NLS--KAAQDTLKNAVKELTSPIVHGNNSGKDGNTSANSADESVKGPNLTEISKKITESN   222
               ||||   ||:|    |:|
Db    273  NADALNTAMTNLKNGIQD------------QNTIKQG-VNFTDADE--------------AKRNAYTN   313

Qy    223  AVVLAVKEIETLLASIDELATKAIGKKIQQNGGLAVEAG--------------HNGTLLAGAY   271
               ::|| |||  ||  ||| |:   |:|             ::|
Db    314  AVTQA----------------------EQILNKAQGPNTSKDG---VETALENVQRAKNELNGNQNVANAK   359

Qy    272  TISKLITQKLDGLKNSEK-----LKEKIENAKK----CSEDFTKKLEGEHAQLGIEN--VTDENA   325
               |:|||        |:|:|   |:||         |:|          ||| |
Db    360  TTAKNALNNLTSINNAQKEALKSQIEGATTVAGVNQVSTTASELNTAMSNLQNGINDEAA   419

Qy    326  KKAILITDAAKDKGAAELEKLFKAVENLAKAAKEMLANSVKE   367
               |:|                ||| |  ||| |  |:
Db    420  TKAALNGTQNLEKAKQHANTAIDGLSHLTNAQKEALKQLVQQ   461
```

Middle column

```
RESULT   5
US-10-072-851-5883
; Sequence 5883, Application US/10072851
; GENERAL INFORMATION:
;  APPLICANT: Carr, Grant J.
;  APPLICANT: Xu, H. Howard
;  APPLICANT: Foulkes, J. Gordon
;  APPLICANT: Zamudio, Carlos
;  APPLICANT: Haselbeck, Robert
;  APPLICANT: Ohlsen, Karl L.
;  APPLICANT: Zyskind, Judith W.
;  APPLICANT: Wall, Daniel
;  APPLICANT: Trawick, John D.
;  APPLICANT: Yamamoto, Robert T.
;  APPLICANT: Roemer, Terry
;  APPLICANT: Jiang, Bo
;  APPLICANT: Boone, Charles
;  APPLICANT: Bussey, Howard
;  TITLE OF INVENTION: Methods for Identifying the Target of a Compound which inhibits c
;  TITLE OF INVENTION: Proliferation
;  FILE REFERENCE: ELITRA.028A
;  CURRENT APPLICATION NUMBER: US/10/072,851
;  CURRENT FILING DATE: 2002-02-08
;  PRIOR APPLICATION NUMBER: 60/267,636
;  PRIOR FILING DATE: 2001-02-09
;  NUMBER OF SEQ ID NOS: 15811
;  SOFTWARE: FastSEQ for Windows Version 4.0
;  SEQ ID NO 5883
;  LENGTH: 837
;  TYPE: PRT
;  ORGANISM: Staphylococcus aureus
US-10-072-851-5883
  Query Match    9.8%;  Score 176;  DB 7;  Length 837;
  Best Local Similarity  23.1%;  Pred. No. 0.00026;
  Matches  93;  Conservative  77;  Mismatches 154;  Indels  78;  Gaps  20;

Qy      4  NNSGKDGNASANSADE---SVKGPNLTEISKKITESNAV--VLAVKE-VETLLASI----   53
               ||:|||      ::   :||||||||||||||||||| ||:||| :||||||
Db    100  NHNLEVAKQHANTAIDGLTSLNGPQKAKLKEQVGQATTLPNVQTVRDNAQTLNTAMKGLR   159

Qy     54  DELATKAIGKKIGNGLEANQSKNTSLLSGAYAISDLIAE----KLNVLKNEELKEKIDT   109
               ||| ||  ||||||||||    |||||::|         :|  |:||| |
Db    160  DSIANEAT-IKAGQNYTDASQNKQTDYNSAVTAAKAIIGQTTSPSMNAQEINQAKDQV-T   217

Qy    110  AKQCSTEFTNKLK------SEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVE   164
               ||| || ||||       |:|||  |:|       |:|  || |       ||| |
Db    218  AKQQALNGQENLRTAQTNAKQHLNGLSDLTD--AQKDAVKRQI---EGATHVNEVTQAQN   272

Qy    165  NLS--KAAQDTLKNAVKELTSPIVHGNNSGKDGNTSANSADESVKGPNLTEISKKITESN   222
               |||   ||| |||||  |:|
Db    273  NADALNTAMTNLKNGIQD-------------------QNTIKQG-VNFTDADE-------AKRNAYTN   313

Qy    223  AVVLAVKEIETLLASIDELATKAIGKKIQQNGGLAVEAG--------------HNGTLLAGAY   271
               ::|  || |                                    :|| |
```

Right column

```
Db    314  AVTQA----------------------EQILNKAQGPNTSKDG---VETALENVQRAKNELNGNQNVANAK   359

Qy    272  TISKLITQKLDGLKNSEK-----LKEKIENAKK----CSEDFTKKLEGEHAQLGIEN--VTDENA   325
               |:|||        |:|:|   |:||         |:|          ||| |
Db    360  TTAKNALNNLTSINNAQKEALKSQIEGATTVAGVNQVSTTASELNTAMSNLQNGINDEAA   419

Db    420  TKAALNGTQNLEKAKQHANTAIDGLSHLTNAQKEALKQLVQQ   461

RESULT   6
US-09-815-242-13080
; Sequence 13080, Application US/09815242
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT: Ohlsen, Karl L.
;  APPLICANT: Zyskind, Judith W.
;  APPLICANT: Wall, Daniel
;  APPLICANT: Trawick, John D.
;  APPLICANT: Carr, Grant J.
;  APPLICANT: Yamamoto, Robert T.
;  APPLICANT: Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION: Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE: 2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
;  SEQ ID NO 13080
;  LENGTH: 875
;  TYPE: PRT
;  ORGANISM: Staphylococcus aureus
US-09-815-242-13080
  Query Match    9.8%;  Score 176;  DB 6;  Length 875;
  Best Local Similarity  23.1%;  Pred. No. 0.00028;
  Matches  93;  Conservative  77;  Mismatches 154;  Indels  78;  Gaps  20;

Qy      4  NNSGKDGNASANSADE---SVKGPNLTEISKKITESNAV--VLAVKE-VETLLASI----   53
               ||:|||      ::   :||||||||||||||||||| ||:||| :||||||
Db    137  NHNLEVAKQHANTAIDGLTSLNGPQKAKLKEQVGQATTLPNVQTVRDNAQTLNTAMKGLR   196

Qy     54  DELATKAIGKKIGNGLEANQSKNTSLLSGAYAISDLIAE----KLNVLKNEELKEKIDT   109
               ||| ||  ||||||||||    |||||::|         :|  |:||| |
Db    197  DSIANEAT-IKAGQNYTDASQNKQTDYNSAVTAAKAIIGQTTSPSMNAQEINQAKDQV-T   254

Qy    110  AKQCSTEFTNKLK------SEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVE   164
               ||| || ||||       |:|||  |:|       |:|  || |       ||| |
Db    255  AKQQALNGQENLRTAQTNAKQHLNGLSDLTD--AQKDAVKRQI---EGATHVNEVTQAQN   309

Qy    165  NLS--KAAQDTLKNAVKELTSPIVHGNNSGKDGNTSANSADESVKGPNLTEISKKITESN   222
               |||   ||| |||||  |:|
Db    310  NADALNTAMTNLKNGIQD-------------------QNTIKQG-VNFTDADE-------AKRNAYTN   350

Qy    223  AVVLAVKEIETLLASIDELATKAIGKKIQQNGGLAVEAG--------------HNGTLLAGAY   271
               ::|  || |                                    :|| |
```

```
Db   351  AVTQA----------EQILNKAQGPNTSKDG---VETALENVQRAKNELNGNQNVANAK   396
Qy   272  TISKLITQKLDGLKNSEK--LKEKIENAKK---CSEDFTKKLEGEHAQLGIEN-VTDENA  325
Db   397  TTAKNALNNLTSINNAQKEALKSOIEGATTVAGVNQVSTTASELNTAMSNLQNGINDEAA  456

Qy   326  KKAILITDAAKDKGAAELEKLFKAVENLAKAAKEMLANSVKE  367
Db   457  TKAALNGTQNLEKAKQHANTAIDGLSHLTNAQKEALKQLVQQ  498
```

RESULT 7
US-10-072-851-13080
Sequence 13080, Application US/10072851
GENERAL INFORMATION:
APPLICANT: Carr, Grant J.
APPLICANT: Xu, H. Howard
APPLICANT: Foulkes, J. Gordon
APPLICANT: Zamudio, Carlos
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Roemer, Terry
APPLICANT: Jiang, Bo
APPLICANT: Boone, Charles
APPLICANT: Bussey, Howard
TITLE OF INVENTION: Methods for Identifying the Target of a Compound which Inhibits c
TITLE OF INVENTION:
FILE REFERENCE: ELITRA.028A
CURRENT APPLICATION NUMBER: US/10/072,851
CURRENT FILING DATE: 2002-02-08
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
NUMBER OF SEQ ID NOS: 15811
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 13080
LENGTH: 875
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-10-072-851-13080

```
Query Match        9.8%;  Score 176;  DB 7;  Length 875;
Best Local Similarity   23.1%;  Pred. No. 0.00028;
...ches 93;  Conservative 77;  Mismatches 154;  Indels 78;  Gaps 20;

Qy     4  NNSGKDGNASANSADE---SVKGPNLTEISKKITESNAV--VLAVKE-VETLLASI----   53
Db   137  NIHLEVAKQNANTAIDGLTSLNGTPQKAKILKEQVGQATTLPNVQTVRDNAQTLNTAMKGLR  196
Qy    54  DSIRNEAT-IKAGQNYTDASQNKQTDYNSAVTAAKAIIGQTTSPSMNAQEINQAKDQV-T   109
Db   197  DELATKAIGKKIGNNGLEANQSKNTSLLSGAYAISDLIAE---KLNVLKNEELKEKIDT    254
Qy   110  AKQCSTEFTNKLK-----SEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVE   164
Db   255  AKQOALNGOENLRTAQTNAKQHLNGLSDLTD--AQKDAVKRQI--EGATHVNEVTQAQN    309
Qy   165  NLS--KAAQDTLKNAVKELTSPIVHGNNSGKDGNTSANSADESVKGPNLTEISKKITESN   222
Db   310  NADALNTAMTNLKNGIQD------QNTIKQG-VNFTDADE------AKRNAYTN        350
Qy   223  AVVLAVKEIETLLASIDELATKAIGKKIQOONGGLAVEAG----------HNGTLLAGAY  271
Db   351  AVTQA--------EQILNKAQGPNTSKDG---VETALENVQRAKNELNGNQNVANAK      396
Qy   272  TISKLITQKLDGLKNSEK---LKEKIENAKK---CSEDFTKKLEGEHAQLGIEN-VTDENA 325
Db   397  TTAKNALNNLTSINNAQKEALKSOIEGATTVAGVNQVSTTASELNTAMSNLQNGINDEAA  456
```

```
Qy   326  KRAILITDAAKDKGAAELEKLFKAVENLAKAAKEMLANSVKE  367
Db   457  TKAALNGTQNLEKAKQHANTAIDGLSHLTNAQKEALKQLVQQ  498
```

RESULT 8
US-09-708-427-15045
Sequence 15045, Application US/0970 8427
GENERAL INFORMATION:
APPLICANT: N. ALEXANDROV et al.
TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
TITLE OF INVENTION: THEREBY
FILE REFERENCE: 2750-1243P
CURRENT APPLICATION NUMBER: US/09/708,427
CURRENT FILING DATE: 2000-11-09
NUMBER OF SEQ ID NOS: 85364
SOFTWARE: Patentin version 3.1
SEQ ID NO 15045
LENGTH: 1304
TYPE: PRT
ORGANISM: Arabidopsis thaliana
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1..1304
NAME/KEY: misc_feature
LOCATION: 1..1304
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: 1..1304
OTHER INFORMATION: Ceres Seq. ID 1828628
US-09-708-427-15045

```
Query Match        9.5%;  Score 170;  DB 6;  Length 1304;
Best Local Similarity   22.8%;  Pred. No. 0.0012;
Matches 91;  Conservative 80;  Mismatches 136;  Indels 92;  Gaps 17;

Qy     4  NNSGKDGNASANSADES----------VKGPNLTEISKKITESNAVVLAVKE----EVET   48
Db    85  NEKLKEALAAQKRAEESFEVEKFRAVELEQAGLEAVQKKDVTSKNELESIRSQHALDISA   144
Qy    49  LLASIDELATKAIGKKIGNNGLEANQSKNTSL--LSGAYAISDLIAEKLNVLKNE----  101
Db   145  LISTEEL-----QRVKHELSMTADAKNKALSHAEEAATKIAEIHAEKAEILASELGRLK   198
Qy   102  ------ELKEKIDTAKQCSTEFTNKLKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAE  155
Db   199  ALLGSKEEKPAIE-----GNEIVSKIKSE----IELLRGELEKVSIILESSLKEQEGLVE  248
Qy   156  LEKL----FRAVENLSKAAQDTLKNAVKELTSPIVHGNNSGKDGNTSANSADESVKG--P  209
Db   249  QLKVDLEAAKMAESCTNSSVEEWKRNVHELEKEV--------EESNRSKSSASESMESVMK 301
Qy   210  NLTEISKKITESNAVVLAVKE--IETLLASIDELAT--KAIGKKIQONGGLAVEAGHNGTL 266
Db   302  QLAELNHVLHETKSDNAAQKEKIELLEKTIEAQRIDLEEYGRQV----------        345
Qy   267  LAGAYTISKLITQKLDGLKNSEKLKEKIENAKKCSEDFTKKLEGEHAQL-GIENVTDENA  325
Db   346  ------CIAKEEASKLENL--VESIKSELEISQ--EEKTRALDNEKAATSNIQNLLDQRT  395
Qy   326  KKAILITDAAKDKGAAELEKLFKAVENLAKAAKEMLANS  364
Db   396  ELSIEL-----ERCKVEEERSKKDMESLITLALQEASTES  429
```

RESULT 9
US-09-708-427-15044
Sequence 15044, Application US/0970 8427
GENERAL INFORMATION:
APPLICANT: N. ALEXANDROV et al.
TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
TITLE OF INVENTION: THEREBY
FILE REFERENCE: 2750-1243P
CURRENT APPLICATION NUMBER: US/09/708,427

; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 15044
; LENGTH: 1313
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..1313
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..1313
; OTHER INFORMATION: Ceres Seq. ID 1828627
US-09-708-427-15044

Query Match   9.5%;   Score 170;   DB 6;   Length 1313;
Best Local Similarity   22.8%;   Pred. No. 0.0012;
Matches   91;   Conservative   80;   Mismatches 136;   Indels   92;   Gaps   17;

Qy   4   NNSGKDGNASANSADES----------------VKGPNLTEISKKITESNAVVLAVK------EVET 48
            |:: ||| ||| :|| |||
Db   94   NEKLKEALAAQKRAEESFEVEKFRAVELEQOACLEAVQKKDVTSKNELESIRSQHALDISA 153

Qy   49   LLASIDELATKAIGKKIGNNGLEANQSKNTSL--LSGAYAISDLIAEKLNVLKNE----- 101
            ||| :||| ||| :|| ||| :|| :|
Db   154   LLSTTEEL-------QRVKHELSMTADAKNKALSHAEEATKIAEIHAEKAEILASELGRLK 207

Qy   102   ----ELKEKIDTAKQCSTEFTNKLRSEHAVLGLDNLTDDDNAQRAILKKHANKDKGAAE 155
            |||| :|:   :|| :|
Db   208   ALLGSKEEKEAIE-----GNEIVSKLKSE---IELLRGELEKVSILESSLKEQEGLVE 257

Qy   156   LEKL-----FKAVENLSKAAQDTLKNAVKELTSPIVHGNNSGKDGNTSANSADESVKG--P 209
            |||| |||| :|: |:   :|
Db   258   QLKVDLEAAKMAESCTNSSVEEWKNKVHELEKEV------EESNRSKSSASESMESVMK 310

Qy   210   NLTEISKKITESNAVVLAVKE-IETLLASIDELAT--KAIGKKIQQNGGLAVEAGHNGTL 266
            |:| :|| :|| |:| ||
Db   311   QLAELNHVLHETKSDNAAQKERIELLEKTIEAQRTDLEEYGROV----------------- 354

Qy   267   LAGAYTISKLITQKLDGLKNSEKLKEKIENAKKCSEDFTKKLEGEHAQL-GIENVTDENA 325
            :|| ||: |: |:|: |: ||
Db   355   ---CIAKEEASKLENL--VESIKSELEISQ---EEKTRALDNEKAATSNIQNLLDQRT 404

Qy   326   KKAILITDAAKDKGAAELEKLFKAVENLAKAAKEMLANS 364
            |:| |:| ||| |:| ||
Db   405   ELSIEL-----ERCKVEEEKSRKDMESLTLALQEBASTES 438

RESULT 10
US-09-815-242-5835
; Sequence 5835, Application US/09815242
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; PRIOR FILING DATE: 2001-02-16
; SEQ ID NO 5835
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-5835

Query Match   9.4%;   Score 169.5;   DB 6;   Length 2434;
Best Local Similarity   23.9%;   Pred. No. 0.0029;
Matches   94;   Conservative   76;   Mismatches 144;   Indels   79;   Gaps   21;

Qy   4   NNSGKDGNASANSADE-----SVKGPNLTEISKKITESNAV--VLAVKE-VETLLASI--- 53
            |:: || ||| ||| ||
Db   702   NHNLEVAKQNANTAIDGLTSLNGPQKAKLKEQVGQATTLPNVQTVRDNAQTLNTAMKGLR 761

Qy   54   DELATKAIGKKIGNNGLEANQSKNTSLLSGAYAISDLIAE---KLNVLKNEELKEKIDT 109
            |:| :| |:|| :|| :|| :|| :| :||
Db   762   DSIANEAT-IKAGQNYTDASQNKQTDYNSAVTAAKAIIGQTTSPSMNAQEINQAKDQV-T 819

Qy   110   AKQCSTEFTNKLK------SEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVE 164
            ||| |: |:| :| ||: |: ||
Db   820   AKQQALNGQENLRTAQTNAKQHLNGLSDLTD--AQKDAVKRQI---EGATHVNEVTQAQN 874

Qy   165   NLS--KAAQDTLKNAVKELTSPIVHGNNSGKDGNTSANSADESVKGPNLTEISKKITESN 222
            ||| ||| :|: |:   :|
Db   875   NADALNTAMTNLKNGIQD------ONTIKQG-VNFTDADE---------AKRNAYTN 915

Qy   223   AVVLAVKEIETLLASIDELATKAIGKKIQQNGGLAVEAG----------HNGTLLAGAY 271
            ||| |:| :| : ||
Db   916   AVTQA----------EQILNKAQGPNTSKDG---VETALENVQRAKNELNGNQNVANAK 961

Qy   272   TISKLITQKLDGLKNSEK--LKEKIENAKK---CSEDFTKKLEGEHAQLGIEN-VTDENA 325
            :| ||: |:| |: |:|: |: ||
Db   962   TTAKNALNNLTSINNAQKEALKSQIEGATTVAGVNQVSTTASELNTAMSNLQNGINDEAA 1021

Qy   326   KKAI-LITDAAKDKGAAELEKLFKAVENLAKAA 357
            |||| |:| ||| |:|
Db   1022   TKAAQKYTDADREKQTAYNDAVTAAKTLLDKTA 1054

RESULT 11
US-10-072-851-5835
; Sequence 5835, Application US/10072851
; GENERAL INFORMATION:
; APPLICANT: Carr, Grant J.
; APPLICANT: Xu, H. Howard
; APPLICANT: Foulkes, J. Gordon
; APPLICANT: Zamudio, Carlos
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Roemer, Terry
; APPLICANT: Jiang, Bo
; APPLICANT: Boone, Charles
; APPLICANT: Bussey, Howard
; TITLE OF INVENTION: Methods for Identifying the Target of a Compound which Inhibit
; TITLE OF INVENTION: Proliferation
; FILE REFERENCE: ELITRA.028A
; CURRENT APPLICATION NUMBER: US/10/072,851
; CURRENT FILING DATE: 2002-02-08

```
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 15811
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5835
;   LENGTH: 2434
;   TYPE: PRT
;   ORGANISM: Staphylococcus aureus
US-10-072-851-5835

Query Match    94; Best Local Similarity   9.4%;   Score 169.5;   DB 7;   Length 2434;
Matches   94; Conservative   76; Mismatches 144; Pred. No. 0.0029;   Indels   79;   Gaps   21;

Qy    4 NNSGKDGNASANSADE---SVKGPNLTEISKKITESNAV--VLAVKE-VETLLASI----   53
        ::: :    :::: :  :  |  |     ||  :   :: ||   ::  :| ::  |
Db  702 NHNLEVAKQNANTAIDGLTSLNGPQKRAKLKEQVGQATTLPNVQTVRDNAQTLNTAMKGLR  761

Qy   54 DELATKAIGKKIGNNGLEANQSKNTSLLSGAYAISDLIAE----KLNVLKNEELKEKIDT   109
        :  |   |    | ::| :| : :   :||  |   |  |:|    ||:|| :|| | |::
Db  762 DSIANEAT-IKAGQNYTDASQNKQTDYNSAVTAAKAIIGQTTSPSMNAQEINQAKDQV-T   819

Qy  110 AKQCSTEFTNKLK------SEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVE  164
        |||     |  |       ::  :|| :|| || | ||   ||     |
Db  820 AKQQALNGQENLRTAQTNAKQHLNGLSDLTD--AQKDAVKRQI---EGATHVNEVTQAQN   874

Qy  165 NLS--KAAQDTLKNAVKELTSPIVHGNNSGKDGNTSANSADESVKGPNLTEISKKITESN  222
        ||   | | | |   :|                          ||   | :   :   ||
Db  875 NADALNTAMTNLKNGIQD------QNTIKQG-VNFTDADE-------AKRNAYTN       915

Qy  223 AVVLAVKEIETLLASIDELATKAIGKKIQQNGGLAVEAG-----------HNGTLLAGAY  271
        || |     |  | |  :::  :                           |  :|
Db  916 AVTQA-------EQILNKAQGPNTSKDG---VETALENVQRAKNELNGNQNVANAK      961

Qy  272 TISKLITQKLDGLKNSEK--LKEKIENAKK---CSEDFTKKLEGEHAQLGIEN-VTDENA  325
        |::  |  ||    |  |  | | ||           || |   |        : | |||
Db  962 TTAKNALNNLTSINNAQKEALKSQIEGATTVAGVNQVSTTASELNTAMSNLQNGINDEAA 1021

Qy  326 KKAI-LITDAAKDKGAAELEKLFKAVENLAKA    357
        ||   |::| |  |      |  | | |  |
Db 1022 TKAAQKYTDADREKQTAYNDAVTAAKTLLDKTA 1054
```

RESULT 12
US-09-815-242-12996
Sequence 12996, Application US/09815242
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
Prokaryotes
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22

```
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12996
;   LENGTH: 6281
;   TYPE: PRT
;   ORGANISM: Staphylococcus aureus
US-09-815-242-12996

Query Match    94; Best Local Similarity   9.4%;   Score 169.5;   DB 6;   Length 6281;
Matches   94; Conservative   76; Mismatches 144; Pred. No. 0.0099;   Indels   79;   Gaps   21;

Qy    4 NNSGKDGNASANSADE---SVKGPNLTEISKKITESNAV--VLAVKE-VETLLASI----   53
        ::: :    :::: :  :  |  |     ||  :   :: ||   ::  :| ::  |
Db 1481 NHNLEVAKQNANTAIDGLTSLNGPQKRAKLKEQVGQATTLPNVQTVRDNAQTLNTAMKGLR 1540

Qy   54 DELATKAIGKKIGNNGLEANQSKNTSLLSGAYAISDLIAE----KLNVLKNEELKEKIDT   109
        :  |   |    | ::| :| : :   :||  |   |  |:|    ||:|| :|| | |::
Db 1541 DSIANEAT-IKAGQNYTDASQNKQTDYNSAVTAAKAIIGQTTSPSMNAQEINQAKDQV-T  1598

Qy  110 AKQCSTEFTNKLK------SEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVE  164
        |||     |  |       ::  :|| :|| || | ||   ||     |
Db 1599 AKQQALNGQENLRTAQTNAKQHLNGLSDLTD--AQKDAVKRQI---EGATHVNEVTQAQN  1653

Qy  165 NLS--KAAQDTLKNAVKELTSPIVHGNNSGKDGNTSANSADESVKGPNLTEISKKITESN  222
        ||   | | | |   :|                          ||   | :   :   ||
Db 1654 NADALNTAMTNLKNGIQD------QNTIKQG-VNFTDADE-------AKRNAYTN      1694

Qy  223 AVVLAVKEIETLLASIDELATKAIGKKIQQNGGLAVEAG-----------HNGTLLAGAY  271
        || |     |  | |  :::  :                           |  :|
Db 1695 AVTQA-------EQILNKAQGPNTSKDG---VETALENVQRAKNELNGNQNVANAK     1740

Qy  272 TISKLITQKLDGLKNSEK--LKEKIENAKK---CSEDFTKKLEGEHAQLGIEN-VTDENA  325
        |::  |  ||    |  |  | | ||           || |   |        : | |||
Db 1741 TTAKNALNNLTSINNAQKEALKSQIEGATTVAGVNQVSTTASELNTAMSNLQNGINDEAA 1800

Qy  326 KKAI-LITDAAKDKGAAELEKLFKAVENLAKA    357
        ||   |::| |  |      |  | | |  |
Db 1801 TKAAQKYTDADREKQTAYNDAVTAAKTLLDKTA 1833
```

RESULT 13
US-10-072-851-12996
Sequence 12996, Application US/10072851
GENERAL INFORMATION:
APPLICANT: Carr, Grant J.
APPLICANT: Xu, H. Howard
APPLICANT: Foulkes, J. Gordon
APPLICANT: Zamudio, Carlos
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Trawick, John D.
APPLICANT: Wall, Daniel
APPLICANT: Yamamoto, Robert T.
APPLICANT: Roemer, Terry
APPLICANT: Jiang, Bo
APPLICANT: Boone, Charles
APPLICANT: Bussey, Howard
TITLE OF INVENTION: Methods for Identifying the Target of a Compound which Inhibit
Proliferation
FILE REFERENCE: ELITRA.028A
CURRENT APPLICATION NUMBER: US/10/072,851
CURRENT FILING DATE: 2002-02-08
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
NUMBER OF SEQ ID NOS: 15811
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 12996
LENGTH: 6281

; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-072-851-12996

Query Match   9.4%;   Score 169.5; DB 7;   Length 6281;
Best Local Similarity   23.9%; Pred. No. 0.0099;
Matches   94; Conservative 76; Mismatches 144; Indels 79; Gaps 21;

Qy   4 NNSGKDGNASANSADE---SVKGPNLTEISKKITESNAV--VLAVKE-VETLLASI---- 53
             : :|: || |:| ||| ::|:||| :|| | | : :|| ||| |
Db   1481 NHNLEVAKQNANTAIDGLTSLNGPQRAKLKEQVGQATTLPNVQTVRDNAQTLNTAMKGLR 1540

Qy   54 DELATKAIGKIGNNGLEANQSKNTSLLSGAYAISDLIAE---KLNVLKNEELKEKIDT 109
            :|: |||||||||||||||||||||||||||| :| |||:|||||||||:|
Db   1541 DSIANEAT-IKAGQNYTDASQNKQTDYNSAVTAAKAIIGQTTSPSMNAQEINQAKDQV-T 1598

Qy   110 AKQCSTEFTNKLK------------SEHAVLGLDLNLTDDNAQRAILKKHANKDKGAAELEKLFKAVE 164
            :|| ||| || |:| || || :|| :| || |:|||||||||||
Db   1599 AKQQALNGQENLRTAQTNAKQHLNGLSDLTD--AQKDAVKRQI---EGATHVNEVTQAQN 1653

Qy   165 NLS---KAAQDTLKNAVKELTSPIVHGNNSGKDGNTSANSADESVKGPNLTEISKKITESN 222
            |: |||||: |||| | ||| | :|| : |
Db   1654 NADALNTAMTNLKNGIQD--------ONTIKQG-VNFTDADE------------AKRNAYTN 1694

Qy   223 AVVLAVKEIETLLASIDELATKAIGKKIQONGGLAVEAG-------------HNGTLLAGAY 271
            :|: | | || || || |||||
Db   1695 AVTQA-----------EQILNKAQGPNTSKDG---VETALENVQRAKNELNGNQNVANAK 1740

Qy   272 TISKLITQKLDGLKNSEK--LKEKIENAKK----CSEDFTKKLEGEHAQLGIEN-VTDENA 325
            |||||:|| ||||:| |:|||:||| ||||||||||||||||||| ::||:|
Db   1741 TTAKNALNLTSINNAQKEALKSQIEGATTVAGVNQVSTTASELNTAMSNLQNGINDEAA 1800

Qy   326 KKAI-LITDAAKDKGAAELEKLFKAVENLAKAA 357
            ||| |:|||:||||||||||||||||||||
Db   1801 TKAAQKYTDADREKQTAYNDAVTAAKTLLDKTA 1833

RESULT 14
US-10-037-182-8
GENERAL INFORMATION:
APPLICANT: Tryggvason, Karl
APPLICANT: Doi, Masayuki
APPLICANT: Thyboll, Jill
TITLE OF INVENTION: Recombinant Laminin 10
FILE REFERENCE: 99-274-F
CURRENT APPLICATION NUMBER: US/10/037,182
CURRENT FILING DATE: 2001-12-21
PRIOR APPLICATION NUMBER: 60/257,449
PRIOR FILING DATE: 2000-12-21
PRIOR APPLICATION NUMBER: 60/279,282
PRIOR FILING DATE: 2001-03-28
NUMBER OF SEQ ID NOS: 36
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 8
LENGTH: 1765
TYPE: PRT
ORGANISM: Homo sapiens
US-10-037-182-8

Query Match   9.4%;   Score 168.5; DB 7;   Length 1765;
Best Local Similarity   24.6%; Pred. No. 0.0023;
Matches 102; Conservative 61; Mismatches 160; Indels 91; Gaps 21;

Qy   2 ACNNSGKDG-NASANSADESVKGPNLTEI-------SKKITESNAVVLAVKEVETLLAS 52
            : : :| | :||||||| ||||||||| |||||||||: ||||:|||||
Db   1384 SCSETECGGPNCRTDEGERKCGGPCGGLVTVAHNAWQKAMDLDQDVLSALAEVEQLSKM 1443

Qy   53 IDELATKAIGKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQ 112
            ||||:||||||||||||| ||||||||||||||||||:|| |||||||||:|||||||||
Db   1444 VSEAKLRAD--------EAKQSAEDILLK----TNATKEKMD-KSNEELRNLIKQIRN 1488

Qy   113 CSTEFTNKLKSEHAVLG------------------LDNLTDDNAQRA------ILKKHANKDKG 152
            ||| || | | :|: | |:| |||||||| ||||||||
Db   1489 FLTQDSADLDSIEAVANEVLKMEMPSTPQQLQNLTEDIRERVESLSQVEVILQHSAADIA 1548

Qy   153 AAELEKLFKAVENLSKAAQDT------------LKNAVKELTSPIVHGNNSGKDGNTSANSADESV 206
            ||||||||||||| | ||| |:| || |||||| | ||| | :|| : |
Db   1549 RAEM--LLEEAKRASKSATDVKVTRADMVKEALEEAEKAQVAAEKAIK------QADEDI 1599

Qy   207 KGPN--LTEISKKITESNAVVL-AVKEIETLLASIDELATKAIGKKIQONGGLAVEAGHN 263
            ||| |:|:|||| |||||:| ||||||||||||||||||||| :| || |||
Db   1600 QGTONLLTSLESEFAASEETLFNASQRISELERNVEELKRRA-----AQNSG---EAEYI 1651

Qy   264 GTLLAGAYTISKL---ITQKLDGLKNSEKLKEKIEN--AKKCSEDFTKKLEGEHAQLGIE 318
            |||||||||||| |||||:||||||:|||||| |||||||||||||||||||||
Db   1652 EKVV---YTVKQSAEDVKKTLDG-ELDEKYK-KVENLIAKKTEESSADARRKAEMLQNEAK 1706

Qy   319 NVTDENAKAILITDAAK-----------------DKGAAELEKLFKAVENLAKAAKAREMLA 362
            |: || | |:| || :|||||||||||||||| | | |:
Db   1707 TLLAQANSKLQLLKDPLERKYEDNQRYLEDK-AQELARLEGEVRSLLKDISQKVA 1759

RESULT 15
US-09-938-275-6
Sequence 6, Application US/0993B275
GENERAL INFORMATION:
APPLICANT: Gerardo Castillo
APPLICANT: Alan Snow
TITLE OF INVENTION: Therapeutic and Diagnostic Applications
TITLE OF INVENTION: of Laminin and Laminin-Derived Protein Fragments
FILE REFERENCE: PROTEO.P03
CURRENT APPLICATION NUMBER: US/09/938,275
CURRENT FILING DATE: 2001-08-16
NUMBER OF SEQ ID NOS: 11
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 6
LENGTH: 1786
TYPE: PRT
ORGANISM: Homo Sapiens
PUBLICATION INFORMATION:
DATABASE ACCESSION NUMBER: Swissprot P07942
DATABASE ENTRY DATE: 1988-08-01
US-09-938-275-6

Query Match   9.4%;   Score 168.5; DB 5;   Length 1786;
Best Local Similarity   24.6%; Pred. No. 0.0023;
Matches 102; Conservative 61; Mismatches 160; Indels 91; Gaps 21;

Qy   2 ACNNSGKDG-NASANSADESVKGPNLTEI-------SKKITESNAVVLAVKEVETLLAS 52
            : : :| | :||||||| ||||||||| |||||||||: ||||:|||||
Db   1405 SCSETECGGPNCRTDEGERKCGGPCGGLVTVAHNAWQKAMDLDQDVLSALAEVEQLSKM 1464

Qy   53 IDELATKAIGKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQ 112
            ||||:||||||||||||| ||||||||||||||||||:|| |||||||||:|||||||||
Db   1465 VSEAKLRAD--------EAKQSAEDILLK----TNATKEKMD-KSNEELRNLIKQIRN 1509

Qy   113 CSTEFTNKLKSEHAVLG------------------LDNLTDDNAQRA------ILKKHANKDKG 152
            ||| || | | :|: | |:| |||||||| ||||||||
Db   1510 FLTQDSADLDSIEAVANEVLKMEMPSTPQQLQNLTEDIRERVESLSQVEVILQHSAADIA 1569

Qy   153 AAELEKLFKAVENLSKAAQDT------------LKNAVKELTSPIVHGNNSGKDGNTSANSADESV 206
            ||||||||||||| | ||| |:| || |||||| | ||| | :|| : |
Db   1570 RAEM--LLEEAKRASKSATDVKVTRADMVKEALEEAEKAQVAAEKAIK------QADEDI 1620

Qy   207 KGPN--LTEISKKITESNAVVL-AVKEIETLLASIDELATKAIGKKIQONGGLAVEAGHN 263
            ||| |:|:|||| |||||:| ||||||||||||||||||||| :| || |||
Db   1621 QGTONLLTSLESEFAASEETLFNASQRISELERNVEELKRRA-----AQNSG---EAEYI 1672

Qy   264 GTLLAGAYTISKL---ITQKLDGLKNSEKLKEKIEN--AKKCSEDFTKKLEGEHAQLGIE 318
            |||||||||||| |||||:||||||:|||||| |||||||||||||||||||||
Db   1673 EKVV---YTVKQSAEDVKKTLDG-ELDEKYK-KVENLIAKKTEESSADARRKAEMLQNEAK 1727

```
Qy    319 NVTDENAKKAILITDAAK---------DKGAAELEKLFKAVENLAKAAKEMLA  362
              : :| | : :      | |: |  |  | : |:      | : :|
Db   1728 TLIAQANSKLQLLKDLERKYEDNQRYLEDK-AQELARLEGEVRSILKDISQKVA  1780
```

Search completed: March 18, 2002, 09:58:33
Job time: 391 sec

```
                              GenCore version 4.5
                        Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on:          March 18, 2002, 10:11:55 ; Search time 39.62 Seconds
                                             (without alignments)
                                             340.552 Million cell updates/sec

Title:           US-09-596-746a-42
Perfect score:   1798
Sequence:        1 MACNNSGKDGNASANSADES.........AVENLAKAAKEMLANSVKEL 368

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters:            100059

.num DB seq length:     0
Maximum DB seq length:  200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	644.5	35.8	210	1	OSC1_BORBU	Q07337 borrelia bu
2	588.5	32.7	212	1	OSC2_BORBU	Q08137 borrelia bu
3	346	19.2	214	1	VM24_BORHE	P32778 borrelia he
4	325.5	18.1	215	1	VM03_BORHE	Q02408 borrelia he
5	175	9.7	1790	1	USO1_YEAST	P25386 saccharomyc
6	173.5	9.6	1167	1	CAGA_HELPJ	Q92lt1 helicobacte
7	168.5	9.4	1786	1	LMB1_HUMAN	P07942 homo sapien
8	167	9.3	1013	1	A6OD_DROME	P91927 drosophila
9	161.5	9.0	944	1	NUF1_YEAST	P33380 saccharomyc
10	158	8.8	1433	1	REST_CHICK	P32380 gallus gall
11	158	8.8	1182	1	CGA2_HELPY	P55746 helicobacte
12	157.5	8.8	483	1	M6_STRPY	P08089 streptococc
13	157	8.7	1251	1	RBP2_PLAVB	Q00799 plasmodium
14	157	8.7	1391	1	MST2_DROHY	Q08696 drosophila
15	155.5	8.6	2017	1	MYSN_DROME	Q99323 drosophila
16	155	8.6	2022	1	ANT1_ONCVO	P21249 onchocerca
17	154.5	8.6	2349	1	TPR_HUMAN	P12270 homo sapien
18	153	8.5	1786	1	LMB1_MOUSE	P48679 mus musculu
19	152	8.5	1639	1	LMG1_DROME	P02469 drosophila
20	151.5	8.4	1427	1	REST_HUMAN	P15215 drosophila
21	151.5	8.4	1966	1	MYSB_CAEEL	P30622 caenorhabdi
22	151.5	8.4	3210	1	CENF_HUMAN	P02566 caenorhabdi
23	150	8.3	1509	1	MYSN_ACACA	P49454 homo sapien
24	149.5	8.3	1978	1	MYHB_CHICK	P05659 acanthamoeb
25	148.5	8.3	564	1	M12_STRPY	P10587 gallus gall
26	148.5	8.3	679	1	YKR9_YEAST	P19401 streptococc
27	148	8.2	492	1	M5_STRPY	P02977 streptococc
28	147	8.2	1186	1	SMC_BACSU	P34237 saccharomyc
29	146.5	8.1	1164	1	BAG_STRAG	P51834 bacillus su
30	146.5	8.1	1189	1	SCI1_CHICK	P27951 streptococc
31	146.5	8.1	1230	1	SMC3_YEAST	Q90988 gallus gall
32	146.5	8.1	1957	1	YD86_SCHPO	P47037 saccharomyc
33	145.5	8.1	705	1	YNP9_CAEEL	Q10411 schizosacch
34	145	8.1	662	1	TLPA_BACSU	P39216 bacillus su
35	145	8.1	1147	1	CGA1_HELPY	P80200 helicobacte
36	145	8.1	1186	1	CGA_HELPY	P55980 helicobacte
37	143.5	8.0	539	1	M24_STRPY	P12379 streptococc
38	143.5	8.0	918	1	YMJB_CAEEL	P34487 caenorhabdi
39	143.5	8.0	1233	1	YFI6_YEAST	P43597 saccharomyc
40	143.5	8.0	1972	1	MYHB_HUMAN	P35749 homo sapien
41	143	8.0	724	1	HMR_HUMAN	O75330 homo sapien
42	142.5	7.9	2116	1	MYS2_DICDI	P08799 dictyosteli
43	142	7.9	1084	1	MYSS_RABIT	P02562 oryctolagus
44	142	7.9	1526	1	MYS2_SCHPO	Q9usi6 schizosacch
45	141.5	7.9	3110	1	LMA2_HUMAN	P24043 homo sapien

ALIGNMENTS

```
RESULT 1
OSC1_BORBU
ID   OSC1_BORBU     STANDARD;     PRT;   210 AA.
AC   Q07337;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   OUTER SURFACE PROTEIN C PRECURSOR (PC).
DE   OSPC OR BBB19.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OG   Plasmid lp54;
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 35210 / B31;
RX   MEDLINE=93268136; PubMed=8098841;
RA   Jauris-Heipke S., Fuchs R., Motz M., Preac-Mursic V., Schwab E.,
RA   Will G., Wilske B.;
RT   "Genetic heterogenity of the genes coding for the outer surface
RT   protein C (OspC) and the flagellin of Borrelia burgdorferi.";
RL   Med. Microbiol. Immunol. 182:37-50(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 35210 / B31;
RX   MEDLINE=93239332; PubMed=8478108;
RA   Wilske B., Preac-Mursic V., Jauris S., Pradel I., Soutschek E.,
RA   Schwab E., Wanner G.;
RT   "Immunological and molecular polymorphisms of OspC, an immunodominant
RT   major outer surface protein of Borrelia burgdorferi.";
RL   Infect. Immun. 61:2182-2191(1993).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 35210 / B31;
RX   MEDLINE=9404163O; PubMed=8225587;
RA   Padula S.J., Sampieri A., Dias F., Szczepanski A., Ryan R.W.;
RT   "Molecular characterization and expression of p23 (OspC) from a North
RT   American strain of Borrelia burgdorferi.";
RL   Infect. Immun. 61:5097-5105(1993).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 35210 / B31;
RX   MEDLINE=96025162; PubMed=7494039;
RA   Fukunaga M., Hamase A.;
RT   "Outer surface protein C gene sequence analysis of Borrelia
RT   burgdorferi sensu lato isolates from Japan.";
RL   J. Clin. Microbiol. 33:2415-2420(1995).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 35210 / B31;
RX   MEDLINE=98065943; PubMed=9403685;
RA   Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA   Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
RA   Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
RA   Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,
RA   van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
```

RA Utterback T., Watthey L., McDonald L., Artiach P., Bowman C.,
RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
RA Smith H.O., Venter J.C.;
RT "Genomic sequence of a Lyme disease spirochaete, Borrelia
RT burgdorferi.";
RL Nature 390;580-586(1997).
CC -!- FUNCTION: NOT KNOWN; MAJOR IMMUNODOMINANT PROTEIN.
CC -!- SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID
CC   ANCHOR.
CC ---
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC ---
DR EMBL; X69596; CAA49306.1; -.
DR EMBL; U01894; AAA16058.1; -.
DR EMBL; D49497; BAA08457.1; -.
DR EMBL; AE000792; AAC66329.1; -.
DR TIGR; BBB19; -.
DR InterPro; IPR001800; Lipoprotein_6.
DR Pfam; PF01441; Lipoprotein_6; 1.
DR ProDom; PD001149; Lipoprotein_6; 1.
DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KW Outer membrane; Lipoprotein; Signal; Plasmid; Antigen;
KW Complete proteome.
FT SIGNAL   1   18
FT CHAIN   19   210
FT LIPID   19   19   N-ACYL DIGLYCERIDE (BY SIMILARITY).
SQ SEQUENCE   210 AA; 22340 MW; 7A4FC978F91777BF CRC64;

Query Match   35.8%; Score 644.5; DB 1; Length 210;
Best Local Similarity   75.6%; Pred. No. 5.3e-23;
Matches 136; Conservative 15; Mismatches 28; Indels 1; Gaps 1;

Qy   189 NNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDELATKAIGK 248
              ||||||||||||||||||||||||||||||||||| || |:| ||:|||||| ||||
Db   20 NNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGK 79

Qy   249 KIQONGGLAVEAGHNGTLLAGAYTISKKITQKLDGLKNSEKLKEKIENAKKCSEDFTKKL 308
              |:|| ||| || | || |||| |:| | :| ||:| || ||:|| ||||| :||
Db   80 KIHQNGLDTENNHNGSLLAGAYAISTLIKOKLDGLKN-EGLKEKIDAAKKCSETFTKKL 138

Qy   309 EGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLAKAAKEMLANSVKEL 368
              : |: ||| |: || :::|||||| |||||||||||||||||||::||:|||||||||
Db   139 REKHTDLGKEGYTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKEL 198

RESULT 2
ID OSC2_BORBU   STANDARD;   PRT;   212 AA.
AC Q08137;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE OUTER SURFACE PROTEIN C PRECURSOR (PC).
GN OSPC.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=PKO;
RX MEDLINE=92219995; PubMed=1560779;
RA Fuchs R., Jauris S., Lottspeich F., Preac-Mursic V., Wilske B.,
RA Soutschek E.;
RT "Molecular analysis and expression of a Borrelia burgdorferi gene
RT encoding a 22 kDa protein (pC) in Escherichia coli.";
RL Mol. Microbiol. 6:503-509(1992).

RN [2]
RP SEQUENCE OF 1-205 FROM N.A.
RC STRAIN=DK26;
RX MEDLINE=94075528; PubMed=8253951;
RA Theisen M., Frederiksen B., Lebech A.M., Vuust J., Hansen K.;
RT "Polymorphism in ospC gene of Borrelia burgdorferi and
RT immunoreactivity of OspC protein as a diagnostic antigen.";
RL J. Clin. Microbiol. 31:2570-2576(1993).
CC -!- FUNCTION: NOT KNOWN; MAJOR IMMUNODOMINANT PROTEIN.
CC -!- SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID
CC   ANCHOR.
CC ---
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC ---
DR EMBL; X62162; CAA44093.1; -.
DR EMBL; X73624; CAA52003.1; -.
DR InterPro; IPR001800; Lipoprotein_6.
DR Pfam; PF01441; Lipoprotein_6; 1.
DR ProDom; PD001149; Lipoprotein_6; 1.
KW Outer membrane; Lipoprotein; Signal; Plasmid; Antigen.
FT SIGNAL   1   18
FT CHAIN   19   212
FT LIPID   19   19   N-ACYL DIGLYCERIDE (BY SIMILARITY).
SQ SEQUENCE   212 AA; 22499 MW; C20C231FBF2E7D4 CRC64;

Query Match   32.7%; Score 588.5; DB 1; Length 212;
Best Local Similarity   68.8%; Pred. No. 1.7e-20;
Matches 130; Conservative 20; Mismatches 36; Indels 3; Gaps 3;

Qy   1 MACNNSGKDG-NASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATK 59
              :||||| | | ||:|||||||||||||||||||||||| |:||| |||:||||:|||||
Db   17 ISCNNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETLVLSIDELAKK 76

Qy   60 AIGKKI-GNNGLEANQSKNTNSLLSGAYAISDLIAEKLNYLKN-EELKEKIDTAKCCSTEF 117
              |||| | ||||| | ||| || || ||| |:|:|||| || | |||||||| ||||||
Db   77 AIGQKIDNNNGLAALNNQNGSLLAGAYAISTLIITEKLSKLKNLEELKTEIAKAKKCSEEF 136

Qy   118 TNKLKSEHAVLGLDNLITDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNA 177
              |||| |||| | |||||||:|| ||:|||| |||||||||||||||||:||| ||||| |
Db   137 TNKLKSGHADLGKQDATDDHAKAAILKHATTDKGAKEFKDLFESVEGLLKAAQVALTNS 196

Qy   178 VKELTSPIV 186
              |||| | :|
Db   197 VKELTSPVV 205

RESULT 3
ID VM24_BORHE   STANDARD;   PRT;   214 AA.
AC P32778;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE VARIABLE MAJOR OUTER MEMBRANE LIPOPROTEIN 24 PRECURSOR.
GN VMP24.
OS Borrelia hermsii.
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=140;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SSP. HS1 SEROTYPE 24;
RX MEDLINE=93133110; PubMed=1484486;
RA Restrepo B.I., Kitten T., Carter C.J., Infante D., Barbour A.G.;
RT "Subtelomeric expression regions of Borrelia hermsii linear plasmids

```
CC   are highly polymorphic.";
RL   Mol. Microbiol. 6:3299-3311(1992).
CC   -!- FUNCTION: SERVES TO AVOID THE HOST IMMUNE RESPONSE BY CHANGING
CC       FROM ONE SURFACE EXPOSED VMP TO ANOTHER.
CC   -!- SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID
CC       ANCHOR.
CC   -!- SIMILARITY: STRONG, TO VMP3.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; L04786; AAA22964.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
DR   PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KW   Outer membrane; Lipoprotein; Signal; Plasmid.
FT   SIGNAL      1     18       PROBABLE.
FT   CHAIN      19    214       VARIABLE MAJOR OUTER MEMBRANE
FT                              LIPOPROTEIN 24.
FT   LIPID      19     19       N-ACYL DIGLYCERIDE (PROBABLE).
SQ   SEQUENCE   214 AA;  22541 MW;  F1583F51O246F7C7 CRC64;
```

```
Query Match             19.2%;  Score 346;  DB 1;  Length 214;
Best Local Similarity   42.1%;  Pred. No. 1.2e-09;
Matches   82;  Conservative   40;  Mismatches   57;  Indels   16;  Gaps    5;

Qy    1  MACNNSGKDGNASANSADESVKGP-----NLTEISKKITESNAVVLAVKEVETLLASIDEL   56
         |:|||||| :| |:   |: | |      |:|:|||||| |||||:|:||||||||||||||
Db   17  MSCNNGGPE-----LKSDEVAKSDGTVLDLAKVSKKIKEASAFAASVKEVETLVKSVDEL   71

Qy   57  ATKAIGKKIGNN--GLEANQSKNTSLLSGAYAISDLIAEKLNVLK-----NEELKEKIDTA  110
         | |||||||||   |||||  |||| ||||||:||||||||||        |:|||| ||
Db   72  A-KAIGKKIKNDGGLDTEAGQNGSLIAGVHSVSVSAVKIKVGALETTSGISNELKTKITEV  130

Qy  111  KQCSTEFTNKLKSEHAVLGLDNLTDDNAQRAILKKIHANKDKGAAAELEKLFKAVENLSKAA  170
         |:| | | |::   || |:| :| ::|| ||||: | ||:|| | || |:|:||:||  |
Db  131  KSKAEAFLNKLKDGHTELGKKDASDDDTKRAIKKDNSDKTGASELEALNTAVDALLKAA   190

Qy  171  QDTLKNAVKELTSPI  185
         :||  |||:|||::
Db  191  EGEVEAAIKELTAPV  205
```

Second block (_BORHE):

```
ID   VM03_BORHE     STANDARD;     PRT;   215 AA.
AC   Q02448;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   01-FEB-1994 (Rel. 28, Last annotation update)
DE   VARIABLE MAJOR OUTER MEMBRANE LIPOPROTEIN 3 PRECURSOR.
GN   VMP3.
OS   Borrelia hermsii.
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=140;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SSP. HS1 SEROTYPE 3;
RX   MEDLINE=93133110; PubMed=1484486;
RA   Restrepo B.I., Kitten T., Carter C.J., Infante D., Barbour A.G.;
RT   "Subtelomeric expression regions of Borrelia hermsii linear plasmids
RT   are highly polymorphic.";
RL   Mol. Microbiol. 6:3299-3311(1992).
CC   -!- FUNCTION: SERVES TO AVOID THE HOST IMMUNE RESPONSE BY CHANGING
CC       FROM ONE SURFACE EXPOSED VMP TO ANOTHER.
```

Middle column (next entry):

```
CC   -!- SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID
CC       ANCHOR.
CC   -!- SIMILARITY: STRONG, TO VMP24.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; L04789; AAA22967.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
DR   PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KW   Outer membrane; Lipoprotein; Signal; Plasmid.
FT   SIGNAL      1     18       PROBABLE.
FT   CHAIN      19    215       VARIABLE MAJOR OUTER MEMBRANE
FT                              LIPOPROTEIN 3.
FT   LIPID      19     19       N-ACYL DIGLYCERIDE (PROBABLE).
SQ   SEQUENCE   215 AA;  23139 MW;  684C74D35F87C771 CRC64;
```

```
Query Match             18.1%;  Score 325.5;  DB 1;  Length 215;
Best Local Similarity   42.3%;  Pred. No. 9.8e-09;
Matches   80;  Conservative   32;  Mismatches   62;  Indels   15;  Gaps    5;

Qy  189  NNSGKDGNTSANSADESVKGP-----NLTEISKKITESNAVVLAVKEIETLLASIDELATK  244
         |:|| |  :  |||||||||      |:|:||||||||||||||||:|||||||||||| |
Db   20  NNGGPE-----LKSDEVAKSDGTVLDLAKISKKIKDASDPAASVKEVHTLVKSIDELA-K   73

Qy  245  AIGKKIQ-ONGGLAVEAGHNGTLLAGAYTISKLITQKLDGLKN-----SEKLKEKIENAKK  299
         ||||||  :| ||   |  || |:||| ||||||| |:||| |       |:|:||||||:
Db   74  AIGKKIKNDNSNFEDENDHNGSLIAGVFQVILTVKAKLTSLEQIIGISDELKTEVGMVKK  133

Qy  300  CSEDFTKKLEGEHAQLGIENVTDENAKKAILITDAAKDKGAAAELEKLFKRAVENLAKAAKE  359
         ||:| | |||| || ||| ||||:|:|||||||| |||:|| | || |:|:|||:||  |
Db  134  ESEAFVTQVKSKHTDLAKEGVTDAHAKSAILVTDGTKDKGAAAELIKLNTAIDELLKAAND  193

Qy  360  MLANSVKEL  368
         :||  |||
Db  194  AVETVIKEL  202
```

Right column (RESULT 5 — USO1_YEAST):

```
RESULT 5
USO1_YEAST
ID   USO1_YEAST     STANDARD;     PRT;   1790 AA.
AC   P25386;
DT   01-MAY-1992 (Rel. 22, Created)
DT   01-MAY-1992 (Rel. 22, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   INTRACELLULAR PROTEIN TRANSPORT PROTEIN USO1.
GN   USO1 OR INT1 OR YDL058W.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=X2180-1A;
RX   MEDLINE=91185402; PubMed=2010462;
RA   Nakajima H., Hirata A., Ogawa Y., Yonehara T., Yoda K.,
RA   Yamasaki M.;
RT   "A cytoskeleton-related gene, uso1, is required for intracellular
RT   protein transport in Saccharomyces cerevisiae.";
RL   J. Cell Biol. 113:245-260(1991).
RN   [2]
RP   SEQUENCE OF 782-1790 FROM N.A.
RA   Hostetter M.K., Herman D.J., Bendel C.M., McClellan M., Tao N.,
RA   Kendrick K.E.;
RL   Submitted (FEB-1993) to the EMBL/GenBank/DDBJ databases.
```

RN   [3]
RP   SEQUENCE OF 1-8 FROM N.A.
RA   Bai Y., Symington L.S.;
RL   Submitted (MAY-1996) to the EMBL/GenBank/DDBJ databases.
-!- FUNCTION: REQUIRED FOR PROTEIN TRANSPORT FROM THE ER TO THE GOLGI
CC   COMPLEX.
-!- SUBCELLULAR LOCATION: CYTOPLASMIC. ASSOCIATED WITH INTRACELLULAR
CC   MEMBRANES. PROBABLY PRESENT ON VESICLES OPERATIONAL BETWEEN THE
CC   ER AND THE GOLGI COMPLEX.
-!- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, COMPOSED
CC   OF AN HEPTAPEPTIDE REPEAT PATTERN CHARACTERISTIC OF ALPHA-HELICAL
CC   COILED COILS. MAY FORM FILAMENTOUS STRUCTURES IN THE CELL.
-!- SIMILARITY: BELONGS TO THE VDP/USO1/YBL047C FAMILY.

EMBL; X54378; CAA38253.1; -.
EMBL; L03188; AAB00143.1; -.
EMBL; U53668; AAB66659.1; -.
PIR; A38455; A38455.
HSSP; P80220; 1DIP.
SGD; S0002216; USO1.
InterPro; IPR002017; Spectrin.
Transport; Protein transport; Golgi stack; Cytoskeleton; Coiled coil.
DOMAIN   1   724   GLOBULAR HEAD.
DOMAIN   725   1790   COILED COIL (POTENTIAL).
DOMAIN   465   487   CHARGED (HYPER-HYDROPHILIC).
DOMAIN   991   1790   DISPENSABLE FOR THE PROTEIN FUNCTION.
DOMAIN   1172   1786   ASP/GLU-RICH (ACIDIC).
CONFLICT   847   847   G -> E (IN REF. 2).
CONFLICT   924   924   E -> K (IN REF. 2).
CONFLICT 1253   1253   I -> V (IN REF. 2).
CONFLICT 1319   1319   V -> I (IN REF. 2).
CONFLICT 1461   1461   N -> S (IN REF. 2).
CONFLICT 1581   1581   G -> S (IN REF. 2).
CONFLICT 1600   1600   I -> V (IN REF. 2).
CONFLICT 1661   1661   R -> V (IN REF. 2).
CONFLICT 1772   1772   D -> DEEDDEE (IN REF. 2).
SEQUENCE 1790 AA; 206424 MW; 6CE2B2166E9FD4818 CRC64;

Query Match   9.7%; Score 175; DB 1; Length 1790;
Best Local Similarity   22.7%; Pred. No. 0.51;
Matches 106; Conservative 66; Mismatches 158; Indels 136; Gaps 19;

Qy   3 CNNSGKDG--------NASANSADESVKGPNLTEISKKI-----------TESNAVVLAVK 44
          |||| || :| | | :|| ||
Db   929 CNNLSKEKEHISKELEVYKSRFQSHDNLV--AKLTEKLKSLANNYKDMQAENESLIKAVE 986

Qy   45 E-------VETLLASIDELATKAIGKKIGNNGLEAN--QSKNTSLLSGAYAISDLIAEK 94
          | |||||||||||||||||||||||||||| |||||||||||||||||||||
Db   987 ESKNESSIQLSNLQNKIDSMSQEKENFQIERGSIEKNIEQLKKT-------ISDLEQTK 1038

Qy   95 LNVLKNEE-----------LKEKIDTAKQCSTEFTNKLKS------EHAVLGLDNL- 133
          |:|:|||| ||:||
Db   1039 EEIISKSDSSKDEYESQISLLKEKLETATTANDENVNKISELTKTREELEAELAAYKNLK 1098

Qy   134 -------TDDNAQRAI---------LKKHANKDKG-------AAELEKLFKAVENLS 167
          : |:| |:|:||||| | ||||||||||||||:|
Db   1099 NELETKLETSEKALKEVKENEEHLKEEKIQLEKEATETKQQLNSLRANLESLEKEHEDLA 1158

Qy   168 ---KRAAQDTLKNAVKELTSPIVHGNNSGKDGNTSANSADESVKGEN------------ 210
             |:|||||
Db   1159 AQLKKYEEQIANKERQYNEEISQLN---DEITSTQQENESIKKKNDELEGEVKAMKSTS 1214

Qy   211 LTEISKKITESNAVVLAVKEIE--------------TLLASIDELATKAIGKKIQONGGLAVEAGH 262
          : :|:|: | || | ||| ||:| :|| || | | |
Db   ...

Db   1215 EEQSNLKKSEIDALNLQIKELKKKNETNEASLLESIKSVESETV--KIKE---LQDECNF 1269
Qy   263 NGTLLAGAYTISKLIITQKLDGLKNSEKLKEKIENAKKCSEDFTKKLEGEHAQLGIENVTD 322
Db   1270 K-------EKEVSELEDKLKASEDKNSKYLELQKESEKIKEELDAKITTELKIQLEKI 1319
Qy   323 ENAKKAILITDAAKDKGAAELEKLFKAVENLAKAKEMLANSVKEL 368
Db   1320 TNLSK-------AKEKSESELSRLKKTSSEERKNAEEQLEKLKNEI 1358

RESULT   6
CAGA_HELPJ
ID   CAGA_HELPJ   STANDARD;   PRT; 1167 AA.
AC   Q9ZLT1;
DT   20-AUG-2001 (Rel. 40, Created)
DT   20-AUG-2001 (Rel. 40, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   (CAG PATHOGENICITY ASSOCIATED IMMUNODOMINANT ANTIGEN (120 KDA PROTEIN)
DE   (CAG PATHOGENICITY ISLAND PROTEIN 26).
GN   CAGA OR CAI OR CAG26 OR JHP0495.
OS   Helicobacter pylori J99 (Campylobacter pylori J99).
OC   Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC   Helicobacter.
OX   NCBI_TaxID=85963;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99120557; PubMed=9923682;
RA   Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
RA   Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
RA   Tummino P.J., Caruso A., Uria-Nickelsen M., Mills S.D., Ives C.,
RA   Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
RA   Trust T.J.;
RT   "Genomic sequence comparison of two unrelated isolates of the human
RT   gastric pathogen Helicobacter pylori.";
RL   Nature 397:176-180(1999).
-!- FUNCTION: MAY BE NECESSARY FOR THE TRANSCRIPTION, FOLDING, EXPORT,
CC   OR FUNCTION OF THE CYTOTOXIN.

DR   EMBL; AE001483; AAD06073.1; -.
KW   Antigen; Complete proteome.
DOMAIN   246   249   POLY-THR.
DOMAIN   882   889   POLY-ASN.
SEQUENCE 1167 AA; 129729 MW; FD5E86B81CEBD0F2 CRC64;

Query Match   9.68; Score 173.5; DB 1; Length 1167;
Best Local Similarity   22.2%; Pred. No. 0.38;
Matches 105; Conservative 81; Mismatches 140; Indels 147; Gaps 25;

Qy   5 NSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKE-----------VETLLASID 54
Db   693 NINKDLKDFSKSFEDEFKNGKN-KDFSKAEETLKALKGSVKDLGINPEWISKVENLNAALN 751

Qy   55 ELATKAIGKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKO-- 112
Db   752 EF---------KNGKNKDFSKVTQ-----AKSDLENSIKDVIINQKITDKVDNLNQAV 795

Qy   113 ---CSTEFTNKLKSEHAVLGLDNLTDDN-AQRAILKKHANKDKGAAELEKLFKAVEN-- 165
Db   796 SVAKATGDFSG---VEQALADLKNFSKEQLAOQOAQKNEDFNTGKNSA----LYQSVNGV 848

Qy   166 -------LSKRAAQDTL----KNAVKELTSPIVHGNNSGKDGNTSANSADESVKGPNLTE 213
Db   849 NGTLVGNGLSKAEATTLSKNFSDIKKELNAKL--GNFNNNNNGLENSTE------PIYTQ 901

```
Qy  214  ISKKITESNAVVLAVKEIETLLASIDELATKRAIGK--KIQONGGLAVEAGHN---GTL--  266
              ::|:|   |: |||:| ::: |:|:||  : || |   :| | |:| | |
Db  902  VAKKV---KAKIDRLDQIASGLGDVGQAASFLLKRHDKVDDLSKVGLSANHEPIYATIDD  958

Qy  267  LAGAYTISK----------LITQKLDGLKNS----------EKLKEK-  293
                  :|:|:|        ::|: |   :|| |         :||:|
Db  959  LGGPFPLKRHDKVDDLSKVGLSREQKLTQKIDNLNQAVSEAKASHFDNLDQMIDKLKDST  1018

Qy  294  ------------IENAKKCSEDFTKKLE----GEHAQL--GIEN--------VTDENAKKAI  329
                       :|:||          :|  ||:|  |:||            ::|
Db  1019 KKNVVNLYVESAKKVPTSLSAKLDNYATNSHTRINSNVKNGTPINEKATGMLTQKNSEWLK  1078

Qy  330  LITD--AAKDKGAAELEKLFKAVENLAKAAKEM------------LANSVKEL  368
              |:|||  :|:|| ||   |:|:||::         ||:||:|
Db  1079 LVNDKIVAHNVGSAPL----SAYDKIGFNQKNMKDYSDSFKFSTRLSNAVKDI  1127
```

RESULT 7

LMB1_HUMAN

ID   LMB1_HUMAN   STANDARD;   PRT;   1786 AA.
DT   P07942;
DT   01-AUG-1988 (Rel. 08, Created)
DT   01-AUG-1988 (Rel. 08, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   LAMININ BETA-1 CHAIN PRECURSOR (LAMININ B1 CHAIN).
GN   LAMB1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87280097; PubMed=3611077;
RA   Pikkarainen T., Eddy R., Fukushima Y., Byers M., Shows T.,
RA   Pihlajaniemi T., Saraste M., Tryggvason K.;
RT   "Human laminin B1 chain. A multidomain protein with gene (LAMB1)
RT   locus in the q22 region of chromosome 7.";
RL   J. Biol. Chem. 262:10454-10462(1987).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90358768; PubMed=1975589;
RA   Vuolteenaho R., Chow L.T., Tryggvason K.;
RT   "Structure of the human laminin B1 chain gene.";
RL   J. Biol. Chem. 265:15611-15616(1990).
RN   [3]
RP   SEQUENCE OF 1276-1709 FROM N.A.
RX   MEDLINE=88021029; PubMed=3661559;
RA   Jaye M., Modi W.S., Ricca G.A., Mudd R., Chiu I.M., O'Brien S.J.,
RA   Drohan W.N.;
RT   "Isolation of a cDNA clone for the human laminin-B1 chain and its
RT   gene localization.";
RL   Am. J. Hum. Genet. 41:605-615(1987).
CC   -!- FUNCTION: BINDING TO CELLS VIA A HIGH AFFINITY RECEPTOR, LAMININ
CC   IS THOUGHT TO MEDIATE THE ATTACHMENT, MIGRATION, & ORGANIZATION OF
CC   CELLS INTO TISSUES DURING EMBRYONIC DEVELOPMENT BY INTERACTING
CC   WITH OTHER EXTRACELLULAR MATRIX COMPONENTS.
CC   -!- SUBUNIT: LAMININ IS A COMPLEX GLYCOPROTEIN, CONSISTING OF THREE
CC   DIFFERENT POLYPEPTIDE CHAINS (ALPHA, BETA, GAMMA), WHICH ARE BOUND
CC   TO EACH OTHER BY DISULFIDE BONDS INTO A CROSS-SHAPED MOLECULE
CC   COMPRISING ONE LONG & THREE SHORT ARMS WITH GLOBULES AT EACH END.
CC   THE BETA-1 CHAIN IS A SUBUNIT OF LAMININ-1 (EHS LAMININ), LAMININ-
CC   2 (MEROSIN), AND LAMININ-6 (K-LAMININ).
CC   -!- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC   -!- TISSUE SPECIFICITY: FOUND IN THE BASEMENT MEMBRANES (MAJOR
CC   COMPONENT).
CC   -!- DOMAIN: THE ALPHA-HELICAL DOMAINS I AND II ARE THOUGHT TO INTERACT
CC   WITH OTHER LAMININ CHAINS TO FORM A COILED COIL STRUCTURE.
CC   -!- DOMAIN: DOMAINS VI AND IV ARE GLOBULAR.
CC   -!- SIMILARITY: CONTAINS 1 LAMININ N-TERMINAL DOMAIN (DOMAIN VI).
CC   -!- SIMILARITY: CONTAINS 12.5 LAMININ EGF-LIKE DOMAINS.
CC   -!- SIMILARITY: CONTAINS 1 LAMININ DOMAIN IV.
CC   --

```
DR   EMBL; M61951; AAA59486.1; -
DR   EMBL; M58147; AAA59486.1; JOINED.
DR   EMBL; M61917; AAA59486.1; JOINED.
DR   EMBL; M61918; AAA59486.1; JOINED.
DR   EMBL; M61921; AAA59486.1; JOINED.
DR   EMBL; M61922; AAA59486.1; JOINED.
DR   EMBL; M61923; AAA59486.1; JOINED.
DR   EMBL; M61924; AAA59486.1; JOINED.
DR   EMBL; M61925; AAA59486.1; JOINED.
DR   EMBL; M61926; AAA59486.1; JOINED.
DR   EMBL; M61927; AAA59486.1; JOINED.
DR   EMBL; M61928; AAA59486.1; JOINED.
DR   EMBL; M61929; AAA59486.1; JOINED.
DR   EMBL; M61930; AAA59486.1; JOINED.
DR   EMBL; M61931; AAA59486.1; JOINED.
DR   EMBL; M61932; AAA59486.1; JOINED.
DR   EMBL; M61933; AAA59486.1; JOINED.
DR   EMBL; M61934; AAA59486.1; JOINED.
DR   EMBL; M61935; AAA59486.1; JOINED.
DR   EMBL; M61936; AAA59486.1; JOINED.
DR   EMBL; M61938; AAA59486.1; JOINED.
DR   EMBL; M61939; AAA59486.1; JOINED.
DR   EMBL; M61940; AAA59486.1; JOINED.
DR   EMBL; M61941; AAA59486.1; JOINED.
DR   EMBL; M61942; AAA59486.1; JOINED.
DR   EMBL; M61943; AAA59486.1; JOINED.
DR   EMBL; M61944; AAA59486.1; JOINED.
DR   EMBL; M61945; AAA59486.1; JOINED.
DR   EMBL; M61946; AAA59486.1; JOINED.
DR   EMBL; M61947; AAA59486.1; JOINED.
DR   EMBL; M61948; AAA59486.1; JOINED.
DR   EMBL; M61949; AAA59486.1; JOINED.
DR   EMBL; M61950; AAA59486.1; JOINED.
DR   EMBL; M55370; AAA59485.1; -
DR   EMBL; M55371; AAA59485.1; JOINED.
DR   EMBL; M55372; AAA59485.1; JOINED.
DR   EMBL; M55373; AAA59485.1; JOINED.
DR   EMBL; M55374; AAA59485.1; JOINED.
DR   EMBL; M55375; AAA59485.1; JOINED.
DR   EMBL; M55376; AAA59485.1; JOINED.
DR   EMBL; M55344; AAA59485.1; JOINED.
DR   EMBL; M55345; AAA59485.1; JOINED.
DR   EMBL; M55346; AAA59485.1; JOINED.
DR   EMBL; M55347; AAA59485.1; JOINED.
DR   EMBL; M55348; AAA59485.1; JOINED.
DR   EMBL; M55349; AAA59485.1; JOINED.
DR   EMBL; M55350; AAA59485.1; JOINED.
DR   EMBL; M55351; AAA59485.1; JOINED.
DR   EMBL; M55352; AAA59485.1; JOINED.
DR   EMBL; M55353; AAA59485.1; JOINED.
DR   EMBL; M55354; AAA59485.1; JOINED.
DR   EMBL; M55355; AAA59485.1; JOINED.
DR   EMBL; M55356; AAA59485.1; JOINED.
DR   EMBL; M55357; AAA59485.1; JOINED.
DR   EMBL; M55358; AAA59485.1; JOINED.
DR   EMBL; M55359; AAA59485.1; JOINED.
DR   EMBL; M55360; AAA59485.1; JOINED.
DR   EMBL; M55361; AAA59485.1; JOINED.
DR   EMBL; M55362; AAA59485.1; JOINED.
DR   EMBL; M55363; AAA59485.1; JOINED.
DR   EMBL; M55364; AAA59485.1; JOINED.
DR   EMBL; M55365; AAA59485.1; JOINED.
DR   EMBL; M55366; AAA59485.1; JOINED.
DR   EMBL; M55367; AAA59485.1; JOINED.
DR   EMBL; M55368; AAA59485.1; JOINED.
```

DR   EMBL; M55369; AAA59485.1; JOINED.
DR   EMBL; M61916; AAA59482.1; -.
DR   EMBL; M20206; AAA59487.1; -.
DR   PIR; S13547; MMHUB1.
DR   HSSP; P01543; 1BHP.
MIM; 150240.
DR   InterPro; IPR0005561; EGF-like.
DR   InterPro; IPR001886; LamNt.
DR   InterPro; IPR0020049; Laminin_EGF.
DR   Pfam; PF00053; laminin_EGF; 13.
DR   Pfam; PF00055; laminin_Nterm; 1.
DR   PRINTS; PR00011; EGFLAMININ.
DR   ProDom; PD000011; LamNt; 1.
DR   ProDom; PD002082; LamNt; 11.
DR   SMART; SM00180; EGF_Lam; 11.
DR   SMART; SM00136; LamNt; 1.
DR   PROSITE; PS00022; EGF_1; 9.
DR   PROSITE; PS01186; EGF_2; 2.
DR   PROSITE; PS01248; LAMININ_TYPE_EGF; 11.
DR   Laminin EGF-like domain; Basement membrane; Extracellular matrix; Coiled coil;
DR   Glycoprotein; Repeat; Signal.
CHAIN   22   1786   LAMININ BETA-1 CHAIN.
SIGNAL   1   21
CHAIN   22   270   LAMININ N-TERMINAL (DOMAIN VI).
DOMAIN   271   540   4.5 X LAMININ EGF-LIKE REPEATS (DOMAIN
                           V).
DOMAIN   271   334   LAMININ EGF-LIKE 1.
DOMAIN   335   397   LAMININ EGF-LIKE 2.
DOMAIN   398   457   LAMININ EGF-LIKE 3.
DOMAIN   458   509   LAMININ EGF-LIKE 4.
DOMAIN   510   540   LAMININ EGF-LIKE 5 (INCOMPLETE).
DOMAIN   541   771   LAMININ DOMAIN IV.
DOMAIN   772   1178   8 X LAMININ EGF-LIKE REPEATS (DOMAIN
                           III).
DOMAIN   773   820   LAMININ EGF-LIKE 6.
DOMAIN   821   866   LAMININ EGF-LIKE 7.
DOMAIN   867   916   LAMININ EGF-LIKE 8.
DOMAIN   917   975   LAMININ EGF-LIKE 9.
DOMAIN   976   1027   LAMININ EGF-LIKE 10.
DOMAIN   1028   1083   LAMININ EGF-LIKE 11.
DOMAIN   1084   1131   LAMININ EGF-LIKE 12.
DOMAIN   1132   1178   LAMININ EGF-LIKE 13.
DOMAIN   1179   1397   DOMAIN II.
DOMAIN   1398   1430   DOMAIN I.
DOMAIN   1431   1786   DOMAIN ALPHA.
DOMAIN   1216   1315   COILED COIL (POTENTIAL).
DOMAIN   1353   1388   COILED COIL (POTENTIAL).
DOMAIN   1442   1781   COILED COIL (POTENTIAL).
DISULFID 271   298   BY SIMILARITY.
DISULFID 273   280   BY SIMILARITY.
DISULFID 300   309   BY SIMILARITY.
DISULFID 312   332   BY SIMILARITY.
DISULFID 335   344   BY SIMILARITY.
DISULFID 337   362   BY SIMILARITY.
DISULFID 365   374   BY SIMILARITY.
DISULFID 377   395   BY SIMILARITY.
DISULFID 398   411   BY SIMILARITY.
DISULFID 400   426   BY SIMILARITY.
DISULFID 428   437   BY SIMILARITY.
DISULFID 440   455   BY SIMILARITY.
DISULFID 458   472   BY SIMILARITY.
DISULFID 460   479   BY SIMILARITY.
DISULFID 481   490   BY SIMILARITY.
DISULFID 493   507   BY SIMILARITY.
DISULFID 773   785   BY SIMILARITY.
DISULFID 775   792   BY SIMILARITY.
DISULFID 794   803   BY SIMILARITY.
DISULFID 806   818   BY SIMILARITY.
DISULFID 821   833   BY SIMILARITY.
DISULFID 823   840   BY SIMILARITY.
DISULFID 842   851   BY SIMILARITY.
DISULFID 854   864   BY SIMILARITY.
DISULFID 867   876   BY SIMILARITY.
DISULFID 869   883   BY SIMILARITY.

FT   DISULFID   886   895   BY SIMILARITY.

Query Match   9.4%;   Score 168.5;   DB 1;   Length 1786;
Best Local Similarity 24.6%;   Pred. No. 0.99;
Matches 102; Conservative 61; Mismatches 160; Indels 91; Gaps 21;

Qy   2 ACNNSGKDG-NASANSADESVKGPNLTEI--------SKKITESNAVVLAVKEVETLLAS 52
          :|:| | ||:|||||||| |||||| |:|: |:|| |||||||||||||
Db   1405 SCSETECGGPNCRTDEGERKCGGPCGGGLVTVAHNAWQKAMDLDQDVLSALAEVEQLSKM 1464

Qy   53 IDELATKAIGKKIGNNGLEANQSKNTSILSGAYAISDLIAEKLNVLKNEELKEKIDTAKQ 112
          ||:| || | |||||||||||||||||||||||||||||||| |||||||
Db   1465 VSEAKLRAD--------EAKQSAEDILLK-----TNATKEKMD-KSNEELRNLIKQIRN 1509

Qy   113 CSTEFTNKLKSEHAVLG----------------------LDNLTDDNAQRA------- 152
          |:|:| ||||||||||||
Db   1510 FLTQDSADIDDSIEAVANEVLKMEMPSTPQQLQNTEDIRERVESLSQVEVILQHSRADIA 1569

Qy   153 AAELEKLFKAVENLSKAAQDT---------LKNAVKELTSPIVHGNNSGKDGNTSANSADESSV 206
          |||||||||||||||||||| |:|||||:|||:|||||||||||||||||||
Db   1570 RAEM---LLEEAKRASKSAFDVKVTADMVKEALEEEAKAQVAAEKAIK------QADEDI 1620

Qy   207 KGPN---LTEISKKITESNAVVL-AVKELETLLASIDELATKAIGKKIQONGGLAVEAGHN 263
          :| ||:| ||:|||||||| |||||||||||||||||||||||||||:|:
Db   1621 QGTQNLLTSIESETAASEETLFNASQRISELERNVEELKRKA------AQNSG--EAEYI 1672

Qy   264 GTLLAGAYTISKL----ITPQKLDGLKNSEKLKEKIEN--AKKCSEDFFKKLEGEHAQLGIE 318
          ||: ||||||||||||||| | |||: |::||||||||| |
Db   1673 EKVV---YTVKQSAEDVKKTLDG-ELDEKYK-KVENILAKKTEESADARRKAEMLQNEAK 1727

Qy   319 NVTDENAKKAIILITDAAK-----------------DKGAAELEKLFKAVENLAKAAKEMLA 362
          ||:: ||||||||||| ||:|||||||||||||||||:: |
Db   1728 TLLAQANSKLQLLIKDLERKYEDNQRYLEDK-AQELARLEGEVRSLLKDISQKVA 1780

RESULT 8
A60D_DROME

ID   A60D_DROME   STANDARD;   PRT;   1013 AA.
AC   P91927; Q9W160;
DT   30-MAY-2000 (Rel. 39, Created)
DT   20-AUG-2001 (Rel. 40, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   CALCIUM-BINDING MITOCHONDRIAL PROTEIN ANON-60DA.
GN   ANON-60DA OR CG4589.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,

```
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [2]
RP   SEQUENCE OF 626-944 FROM N.A.
RC   TISSUE=Ovary;
RX   MEDLINE=99168769; PubMed=10071211;
RA   Caggese C., Ragone G., Perrini B., Moschetti R., de Pinto V.,
RA   Caizzi R., Barsanti P.;
RT   "Identification of nuclear genes encoding mitochondrial proteins:
RT   isolation of a collection of D. melanogaster cDNAs homologous to
RT   sequences in the Human Gene Index database.";
RL   Mol. Gen. Genet. 261:64-70(1999).
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL.
CC   -!- SIMILARITY: CONTAINS 2 EF-HAND CALCIUM-BINDING DOMAINS.
CC   -!- CAUTION: REF 2 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO
CC       FRAMESHIFTS IN POSITIONS 920 AND 930.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; AE003464; AAF47217.1; -.
DR   EMBL; Y10912; CAA71853.1; ALT_FRAME.
DR   FlyBase; FBgn0019886; CG4558.
DR   InterPro; IPR001813; 60s_ribosomal.
DR   InterPro; IPR002024; Bacterioferritin.
DR   InterPro; IPR003871; DUF223.
DR   InterPro; IPR002048; EF-hand.
DR   InterPro; IPR000861; REM_repeat.
DR   InterPro; IPR001236; ldh.
DR   InterPro; IPR002555; RF1.
DR   Pfam; PF00428; 60s_ribosomal; 1.
DR   Pfam; PF01334; Bacterioferr; 1.
DR   Pfam; PF02721; DUF223; 1.
DR   Pfam; PF00036; efhand; 2.
DR   Pfam; PF02185; HR1; 1.
DR   Pfam; PF00056; ldh; 1.
DR   Pfam; PF01605; RF1; 1.
DR   SMART; SM00054; EFh; 2.
DR   PROSITE; PS00018; EF_HAND; 2.
KW   Mitochondrion; Calcium-binding; Repeat.
FT   CA_BIND       700    711       EF-HAND 1 (POTENTIAL).
FT   CA_BIND       783    794       EF-HAND 2 (POTENTIAL).
FT   CONFLICT      687    687       E -> Q (IN REF. 2).
FT   CONFLICT      690    690       E -> G (IN REF. 2).
FT   CONFLICT      717    717       Q -> H (IN REF. 2).
FT   CONFLICT      736    736       E -> D (IN REF. 2).
FT   CONFLICT      740    740       K -> R (IN REF. 2).
FT   CONFLICT      761    761       K -> E (IN REF. 2).
FT   CONFLICT      943    943       G -> S (IN REF. 2).
SQ   SEQUENCE   1013 AA;  113579 MW;  8C71038026 3F262E CRC64;
```

```
RESULT   9
ID   NUF1_YEAST     STANDARD;      PRT;    944 AA.
AC   P32380;
DT   01-OCT-1993 (Rel. 27, Created)
DT   01-OCT-1993 (Rel. 27, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   NUF1 PROTEIN (SPINDLE POLY BODY SPACER PROTEIN SPC110).
GN   NUF1 OR SPC110 OR YDR356W OR D9476.3.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288C;
RX   MEDLINE=92176232; PubMed=1541631;
RA   Mirzayan C., Copeland C.S., Snyder M.;
RT   "A spacer protein in the Saccharomyces cerevisiae spindle poly body
RT   whose transcript is cell cycle-regulated.";
RL   J. Cell Biol. 116:1319-1332(1992).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288C;
RX   MEDLINE=94064779; PubMed=7503995;
RA   Kilmartin J.V., Dyos S.L., Kershaw D., Finch J.T.;
RT   "The NUF1 gene encodes an essential coiled-coil related protein that
RT   is a potential component of the yeast nucleoskeleton.";
RL   J. Cell Biol. 123:1175-1184(1993).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288C / AB972;
RA   Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z.,
RA   Favello A., Fulton L., Gattung S., Greco T., Kirsten J.,
RA   Kucaba T., Hallsworth K., Hawkins J., Hillier L., Jier M.,
RA   Johnson D., Johnston L., Langston Y., Latreille P., Le T.,
RA   Mardis E., Menezes S., Miller N., Mills N., Pauley K., Peluso D.,
RA   Rifken L., Riles L., Taich A., Trevaskis E., Vignati D.,
RA   Wilcox L., Wohldman P., Vaudin M., Wilson R., Waterston R.;
RL   Submitted (JUN-1995) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: ESSENTIAL COMPONENT OF THE NUCLEOSKELETON. POTENTIAL
CC       ROLE IN CROSSLINKING FILAMENTS OR ANCHORING OTHER MOLECULES. IT
CC       IS ESSENTIAL FOR GROWTH.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR; TIGHTLY ASSOCIATED WITH THE
CC       NUCLEUS. IT IS PRESENT IN A GRANULAR PATTERN THAT EXCLUDES THE
```

```
Query Match         9.38;  Score  167.5;  DB 1;  Length 1013;
Best Local Similarity   22.4%;  Pred. No. 0.6;
Matches   77;  Conservative   63;  Mismatches  147;  Indels   57;  Gaps  11;

Qy   29  ISKKITESNAVVLAVKEVETLLASIDELATKAIGKKIIGNNGLEANQSKNTSLLSGAYAIS  88
                  ::  |||  |  :||  :|:    :|:|
Db  609  VKEPVRESRAAKLLYNRVNKMISQLDNVL------------NDLEARQHQIKQAESSDYAAS  658

Qy   89  DLIAEKLNVLKNEELKEKIDFAKQCSTEFTNKLKSEHAVLGLDNLITDDNAQRAILKKHAN  148
                  ::  |||  |  |  |  |:    |  |:  ||
Db  659  SPTVEPQOMVHIDELVATIRRMKEASDEERFKVVGD------------------LLVKLDAD  702

Qy  149  KDKGAAELEKLFKAVENLSKAAQDTLKNAVKELT----------------SPIVHGNNSGKDGN  196
                   |  ||  ::||  |  |  :|  |   |::             |  |  ||||
Db  703  KD-GVISVNEITKAVQSIDREATNIDKKQLEEFTELLSKLASRRHEEIVHIDDLMNNIK  761

Qy  197  TSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDELATKAIGKKIQQNGGL  256
                    :|:  |:   |  :  |:|    |   |||
Db  762  VLKEITSDEA-RLKHIEAVLEKFDADKDGVVTVNDIRKVLESIGRDNIKLSDKAIEE---  816

Qy  257  AVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKEIENAKKCCSEDFTKLEGEH-AQL  315
                   |  |  ||  |  |  |:  |    |||
Db  817  LISLLDKEQVLQAEQKIEKAIAK---SMKEAEKLKSEVDKADK---DLSKLVNDIHDSAK  870

Qy  316  GIENYTDENAKKAILITDAAKDKGAAELEKLFKAVENLAKAAKE  359
                   |  |    |  |:
Db  871  EIQDIANEMRDKEEETYPDKAKELKA---EPAFK---DTAKTLKD  908
```

```
CC   -!- PTM: MAY BE REGULATED BY PHOSPHORYLATION EVENTS.
CC
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
```

```
DR   EMBL; Z11582; CAA77668.1; -.
DR   EMBL; X73297; CAA51733.1; -.
DR   EMBL; U28372; AAB64791.1; -.
DR   PIR; S26710; S26710.
DR   PIR; S34288; S34288.
DR   SGD; S000276J; NUF1.
```

Coiled coil; Nuclear protein; Phosphorylation.

```
FT   DOMAIN        54     59       COILED COIL.
FT   DOMAIN       726    731       NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT   DOMAIN       742    747       NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT   DOMAIN       731    944       ARG/TYR-RICH.
FT   DOMAIN       164    791       NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
SQ   SEQUENCE    944 AA;  111781 MW;  04FAA074BB8A0BC8 CRC64;
```

Query Match   9.0%;   Score 161.5;   DB 1;   Length 944;
Best Local Similarity   23.2%;   Pred. No. 1;
Matches   100;   Conservative   71;   Mismatches   167;   Indels   93;   Gaps   21;

```
Qy      6 SGKDGNASANSADESVKGPNLTEISKKITESNAV-VLAVKEVETLLASIDE-LATKAIGK 63
          | ||   |         |   :| |||    || |   || |||||| :|   | ||| ||
Db     91 SPRSGN----VDKSRKRNLIDDLKRDVPMSQPLKEQEVREHQMKKERPDRALESKLLGK 145
```

```
Qy     64 K---------IGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKN-EELKEKIDTAKQCS 114
            :       | :     |   || :::|   | |  :|| ::| ::|||||||||| |
Db    146 RHITYANSDISNKELYINEIKSLK------HEIKELRKEKNDTLLNNYDTLEEETDDLKNRL 200
```

```
Qy    115 TEFTNKLKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAV-------ENLS 167
           | :  |  || : |:  :|| :||  |    |:  ||||:||||||||       |||
Db    201 QALEKELDARNKIVN-SRKVDDHS--GCIEEREQMERKLAELERKLKTVKDQVLELENNS 257
```

```
Qy    168 ------KAAQDTLKNAVKELTSPIVHGNNSGKDGNTS------------ANSADE-SVKGPN 210
               | ||| || |    :| | | |   |             | :|   :  :||
Db    258 DVQSLKLRSKEDELKNLMNELNE--LKSNAEEKDTQLEFKKNELRKRTNELNELKIKSDE 315
```

```
Qy    211 L-TEISKKITESNAVVLAVKEIEILL------ASIDELATKAIGKKIQQ-------NGG 255
          | :||  |  |  ::|   |    |       ||||| | |||| |        |||
Db    316 MDLQLKQKQNESKRLKDELNELETKFSENGSQSSAKENELKMLKNKIAELEEEISTKNSQ 375
```

```
Qy    256 LAVEAGHNGTLLAGAYTISKLITQKLDGL----KNSEKLKEKIENAKKCSF----- 302
          |   |    |    |   | :|| |   |    :|::||:|  |:| | |
Db    376 LIAKEGKLASLMAQLTQLESKLNQRDSQLGSREEELKKTKNDKLQKDIRIAREETVSKDER 435
```

```
Qy    303 --DFTKLEGEHAQLGIENVTDENAKKALLITDAAKDKGAAELE---KLFKAVENLAKAA 357
               :|| :|  :|  |  |::|   |::   || | |  ||   |||||||| | ||
Db    436 IIDLQKKVKQLENDLFVIKTHSESK---TITD----NELESKDKLIKILENDLKVA 485
```

```
Qy    358 KEMLANSVKEL 368
          :| |||  ||
Db    486 QEKYSKMEKEL 496
```

```
OS   Gallus gallus (Chicken).
```

```
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98137792; PubMed=9469933;
RA   Griparic L., Volosky J.M., Keller T.C. III;
RT   "Cloning and expression of chicken CLIP-170 and restin isoforms.";
RL   Gene 206:195-208(1998).
RN   [2]
RP   SEQUENCE OF 17-1139 FROM N.A. (CLIP-170(11) AND CLIP-170(11+35)).
RC   TISSUE=Pectoralis muscle;
RA   Griparic L., Keller T.C. III;
RT   "Identification and expression of two novel CLIP-170/Restin isoforms
RT   expressed predominantly in muscle.";
RL   Submitted (FEB-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: SEEMS TO BE A INTERMEDIATE FILAMENT ASSOCIATED PROTEIN
CC       THAT LINKS ENDOCYTIC VESICLES TO MICROTUBULES (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC. ASSOCIATED WITH THE
CC       CYTOSKELETON (BY SIMILARITY).
CC   -!- ALTERNATIVE PRODUCTS: AT LEAST 4 ISOFORMS ARE PRODUCED BY
CC       ALTERNATIVE SPLICING.
CC   -!- SIMILARITY: CONTAINS 2 CAP-GLY DOMAINS.
CC
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
```

```
DR   EMBL; AF014012; AAC60344.1; -.
DR   EMBL; AF020764; AAC60345.1; -.
DR   EMBL; AF045650; AAC03547.1; -.
DR   EMBL; AF045651; AAC03548.1; -.
DR   InterPro; IPR000938; CAP-GLY.
DR   InterPro; IPR001878; Znf_CCHC.
DR   Pfam; PF01302; CAP_GLY; 2.
DR   SMART; SM00343; ZnF_C2HC; 1.
DR   PROSITE; PS00845; CAP_GLY_1; 2.
KW   Cytoskeleton; Microtubules; Coiled coil; Alternative splicing.
FT   DOMAIN        79    121       CAP-GLY 1.
FT   DOMAIN       144    207       SER-RICH.
FT   DOMAIN       235    277       CAP-GLY 2.
FT   DOMAIN       305    332       SER-RICH.
FT   DOMAIN       351   1353       COILED COIL (POTENTIAL).
FT   DOMAIN      1414   1427       CCHC-BOX.
FT   VARSPLIC     458    458       MISSING (IN SHORT ISOFORM).
FT   VARSPLIC     458    492       TQTKLEHARIRELEQSLLFRKTKADRLQRELEDTR ->
FT                                 RKRQISEDPEN (IN ISOFORM CLIP-170(11)).
FT   VARSPLIC     458    803       S -> GGSSKVS (IN ISOFORM CLIP-170(11)).
FT   VARSPLIC     458    803       T -> RKRQISEDPENT (IN ISOFORM CLIP-
FT                                 170(11+35)).
FT   CONFLICT     309    309       K -> R (IN REF. 2; AAC03547).
FT   CONFLICT     440    440       E -> V (IN REF. 2; AAC03548).
SQ   SEQUENCE   1433 AA;  161026 MW;  5631CE8683498E23 CRC64;
```

Query Match   8.8%;   Score 158.5;   DB 1;   Length 1433;
Best Local Similarity   22.5%;   Pred. No. 2.2;
Matches   86;   Conservative   78;   Mismatches   152;   Indels   67;   Gaps   16;

```
Qy     18 DESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKAIGKKIGNNGLEANQSKN 77
          ||  | |||||  |  || ::  |  |:|:    ||||  | ||:|||| :|
Db    508 DLAIRVKEVAELRGRLESSK----HIDDVDTSLSLLQEIS--SLQEKMAAAGKE-HQREM 560
```

```
Qy     78 TSLLSGAYAISDLIAEKLNVL-----------KNEELKEKIDTAKQCSTEFTN--KLKSEHA 126
           ::| |||||||||||||| |            |||:|:||||| ||:|| :| |||||
Db    561 SSLKEKFESSEEALRKEIKTLSASNERMGKENESLKTKLDHANKENSDVIELWKSKLESA 620
```

```
Qy    127 VIGLDNLTDDNAQRAILKKHANKDGA--AELEKLFKAVENLSKAAQDTLKN------- 176
```

RESULT 11

```
CGA2_HELPY      STANDARD;      PRT;   1182 AA.
AC   P55746;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   CYTOTOXICITY ASSOCIATED IMMUNODOMINANT ANTIGEN (120 KDA PROTEIN).
GN   CAGA OR CAI.
OS   Helicobacter pylori (Campylobacter pylori).
OC   Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC   Helicobacter.
OX   NCBI_TaxID=210;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 53726 / 84-183;
RX   MEDLINE=92339281; PubMed=8478069;
RA   Tummuru M.K.R., Cover T.L., Blaser M.J.;
RT   "Cloning and expression of a high-molecular-mass major antigen of
RT   Helicobacter pylori: evidence of linkage to cytotoxin production.";
RL   Infect. Immun. 61:1799-1809(1993).
CC   -!- FUNCTION: MAY BE NECESSARY FOR THE TRANSCRIPTION, FOLDING, EXPORT,
CC       OR FUNCTION OF THE CYTOTOXIN.
```

```
DR   EMBL; L11714; -; NOT_ANNOTATED_CDS.
DR   HSSP; P02937; 1MLP.
KW   Antigen.
FT   DOMAIN      878    885       POLY-ASN.
SQ   SEQUENCE   1182 AA;  131503 MW;  C916817E2E57BB4 CRC64;

Query Match             8.8%;  Score 158;  DB 1;  Length 1182;
Best Local Similarity  23.8%;  Pred. No. 1.9;
Matches  101;  Conservative  61;  Mismatches 140;  Indels 122;  Gaps 20;

Qy     24 PNLTEIS-----------KKITESNAVVLAVKEVETLLASIDELATKAIGKKIGNNGLEA 72
          ::||                 ||| ::|| :| |||| :||| |||
Db    541 PNLNLAITSVVRRDLEDKLIAKGLSPQEANKLVKDFLSSNKELVGKAL-----NFNKAV 595

Qy     73 NQSKNTSLSGAY------AISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNKLKSEHAV 127
          ::||       |:||| :|||              ||| :| |||| :||
Db    596 AEAKNT----GNYDEVKRAQKDL--EK--SLKKREHLEKGDVAKNLESKSGNKNKMEAK- 646

Qy    128 LGLDNLTDDNAQR----AILKKHANKD-----KG----AAELEKLFKAVENLSK 168
          :|:|              |:| ||||          ||       :|| || ||
Db    596 AEAKNT----GNYDEVKRAQKDL--EK--SLKKREHLEKGDVAKNLESKSGNKNKMEAK- 646
```

Middle alignment block (RESULT 11):

```
Db    621 IASHQQAMEE---------LKVSFNKGVGAQTAEFAELKTQMERVKLDYENEMSNLKLKQENE 674
             |:|:|       :|:|                    :|:| |:|   :|:
Qy    177 ------AVKELTSPIVHGNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKE 230
Db    675 KSQHLKEIEALKAKLLEVTEEKEQTLENLKAKLSESVEDQHLVEMEDTLNKLQEAEIKVKE 734
Qy    231 IETLLASIDELATKAIGKKIQQNGGLAVEAGHNGTL---LAGAYTISKLITQKL-DGLK 285
             :| |    ||   | | :|  || | : : |:|    |||| ||||| :| |||
Db    735 LDVLQAKCNE-QTKLIGSLTQQ-----IRASEEKLLDLAALQKANSEGKLEIQKLSEQLQ 788
Qy    286 NSEKLKEIENAKKCSEDFTKKLEGEHAQLGIENVTDENAKKAILITDAAKDKGAAEL-- 343
          ::| |                                            |:|  |  |
Db    789 AAEKQIQNLETEK--VSNLTKELQGKEQKL------LDLEKNLSAVNQVKDSLEKEKLQL 839
Qy    344 --EKLFKAV---ENLAKAAKEML 361
            :|| ||:  ||
Db    840 LKEKFTSAVDGAENAQRAMQETI 862
```

Right alignment block (RESULT 11):

```
Db    647 ---------AQANSQKDEIFALINKEANRDARAIAYAQNLKGIKRELSDKLENINKDLKDFSK 700
              |:|                                                  :|:
Qy    169 A-------------AQDT---LKNAVKEL-----------TSPIVHGNNSGKDG 195
Db    701 SFDGFKNGKNKDFSKAEETLKALKGSVKDLGINPEWISKVENLNAALNEFKGNKNKDFSK 760
Qy    196 NTSANSADE-SVKGPNLTE-ISKKITESNAVVLAVK-----EIETLLASIDELATKAIG 247
Db    761 VTQAKSDQENSIKDVIINQKITDKVDELNQAVSVAKIACDFSGVEQALADLKNFSKEQLA 820
Qy    248 KKIQQNGGL------AVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKEKIENAK 298
Db    821 QQAQKNESFNVGKSEIYQSVKNGVNGTLVGNG-------LSGIEATALAKNFSDIKK 870
Qy    299 KCSEDFTKKLEGEHAQLGIEN-----VTDENAKKAILITDAAKDRGAAELEKLFKAVENL 353
Db    871 ELNEKF--KNFNNNNNNGLKNGGEPIYAQVNKKKTGQVASPEEPIYAQVAKKVTKKIDQL 928
Qy    354 AKAA 357
Db    929 NQAA 932
```

RESULT 12

```
M6_STRPY        STANDARD;      PRT;   483 AA.
ID   M6_STRPY
AC   P08089;
DT   01-AUG-1988 (Rel. 08, Created)
DT   01-AUG-1988 (Rel. 08, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   M PROTEIN, SEROTYPE 6 PRECURSOR.
GN   EMM6.
OS   Streptococcus pyogenes.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC   Streptococcus.
OX   NCBI_TaxID=1314;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86118235; PubMed=3511046;
RA   Hollingshead S.K., Fischetti V.F., Scott J.R.;
RT   "Complete nucleotide sequence of type 6 M protein of the group A
RT   Streptococcus. Repetitive structure and membrane anchor.";
RL   J. Biol. Chem. 261:1677-1686(1986).
RN   [2]
RP   SEQUENCE OF 43-122 FROM N.A.
RX   MEDLINE=85166224; PubMed=3885219;
RA   Scott J.R., Pulliam W.M., Hollingshead S.K., Fischetti V.A.;
RT   "Relationship of M protein genes in group A streptococci.";
RL   Proc. Natl. Acad. Sci. U.S.A. 82:1822-1826(1985).
CC   -!- FUNCTION: THIS PROTEIN IS ONE OF THE DIFFERENT ANTIGENIC SEROTYPES
CC       OF PROTEIN M. PROTEIN M IS CLOSELY ASSOCIATED WITH VIRULENCE OF
CC       THE BACTERIUM AND CAN RENDER THE ORGANISM RESISTANT TO
CC       PHAGOCYTOSIS.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. CELL WALL.
CC   -!- SIMILARITY: TO OTHER M PROTEINS.
CC   -!- SIMILARITY: TO OTHER STREPTOCOCCAL AND STAPHYLOCOCCAL PROTEINS
CC       IN THE REGION OF THE MEMBRANE ANCHOR.
```

```
DR   EMBL; M11338; AAA26920.1; -.
DR   PIR; A26297; A26297.
DR   InterPro; IPR001899; Gram_pos_anchor.
DR   InterPro; IPR003345; M_repeat.
DR   Pfam; PF00746; Gram_pos_anchor; 1.
DR   Pfam; PF02370; M; 9.
DR   PRINTS; PR00015; GPOSANCHOR.
```

```
DR   PROSITE; PS00343; GRAM_POS_ANCHORING; 1.
KW   Virulence; Phagocytosis; Cell wall; Duplication; Repeat; Antigen;
KW   Transmembrane; Coiled coil; Signal.
FT   SIGNAL        1      42
FT   CHAIN        43     483       M PROTEIN, SEROTYPE 6.
FT   DOMAIN       43     57        EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    458     477       MEMBRANE ANCHOR.
FT   TRANSMEM    478     483       M PROTEIN, SEROTYPE 6.
FT   DOMAIN       69     138       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      157     269       10 X 7 AA TANDEM REPEATS.
FT   DOMAIN      279     347       4.5 X 25 AA TANDEM REPEATS.
FT   DOMAIN      348     411       TWO DIRECTLY REPEATED 27 AMINO ACID
FT                                 BLOCKS SEPARATED BY 15 AMINO ACIDS.
FT   DOMAIN      412     448       HYDROPHILIC.
FT   DOMAIN      449     454       GLY/PRO-RICH (CELL WALL-SPANNING).
FT                                 CONSERVED IN GRAM-POSITIVE COCCI SURFACE
FT                                 PROTEINS.
SQ   SEQUENCE   483 AA;  53472 MW;  68F87F28DB53A448 CRC64;

Query Match    8.8%;  Score 157.5;  DB 1;  Length 483;
Best Local Similarity   24.8%;  Pred. No. 0.75;  Mismatches  150;  Indels  113;  Gaps  24;
Matches 109;  Conservative   68;

Qy    1  MACNNSG----------KDGNASANSADESVKGP----NLTEISKKITESNAV--------V  40
            ||:||                ||:| ||| | |        ||||||||||||       |
Db    1  MAKNNTNRHYSLRKLKKGTASVAVA-LSVIGAGLVVNTNEVSARVPPRGTVENPDKAREL--V  59

Qy   41  LAVKEVET--LLASIDELATKAIGKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVL  98
            : | ||    |||||                           : | | || :  ||
Db   60  LNKYDVENSMLQAANDKLTTE------NNNL-TDQNKN------------LTTENKNLT  99

Qy   99  ---KN--EELKEKIDTAKQCSTEFTNKLKSEHAVLGLDNLTDDNAQRAILKKHANKDKGA 153
             |:           |  |     |       : |               |
Db  100  DQNKNLTTENKNLTDQNKNLTTE-----NKELKAEE------KGLTKKLSEAEEEA 151

Qy  154  AELEKLFK-AVENLSKAAQDTLKNAV-KELTS------------PIVHGNNSGKD 194
           | :  |  | |  |   |                            || ||||
Db  152  ANKERENKEAIGTLKKTLDETVKDKIAKEQESKETIGTLKKTLDETVKDKIAKEQESKET 211

Qy  195  GNTSANSADESVKGPNLTEISKKITESNAVVLAVKEI----ETLLASI--------- 238
           |    ||     |                  ||            |
Db  212  IGTLKKTLDETVK----DKIAKE-QESKETIGTLKKTLDETVKDKIAREQKSKQDIGALK 266

Qy  239  DELATKAIGKKIIQQNGGLAVEAGHNGTLLAGAYIISKL--ITQKLDGLKNSEK------ 289
            ||    | |              |  |                :  |
Db  267  QELAKKDEGNKVSEASRKGLRRDLDASREAKKQVEKDLANLTAELDKVKEERKQISDASRQ 326

Qy  290  -LKEKIENAKKCSEDFTKKLEGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFK 348
            |:  :| ||          |                    :  |
Db  327  GLRRDLDASREAKKQVEKAILEEANSKLAALEKLNKELEEESKKLIEKEK----AELQAKLE 382

Qy  349  AVENLAKAAKEMLANSVKEL 368
          | : |       |  ::
Db  383  AE--AKALKEQLAKQAEEL 399
```

RESULT 13
```
ID   RBP2_PLAVB     STANDARD;      PRT;   1251 AA.
AC   Q00799;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-APR-1993 (Rel. 25, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   RETICULOCYTE BINDING PROTEIN 2 (FRAGMENT).
GN   RBP2.
OS   Plasmodium vivax (strain Belem).
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=31273;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92315338; PubMed=1617731;
RA   Galinski M.R., Medina C.C., Ingravallo P., Barnwell J.W.;
RT   "A reticulocyte-binding protein complex of Plasmodium vivax
```

```
RT   merozoites ";
RL   Cell 69:1213-1226(1992).
CC   -!- FUNCTION: INVOLVED IN RETICULOCYTE ADHESION SPECIFICALLY BINDS TO
CC       HUMAN RETICULOCYTE CELLS.
CC   -!- SUBCELLULAR LOCATION: MEMBRANE-BOUND (PROBABLE).
CC   --------------------------------------------------------------------
CC       This SWISS-PROT entry is copyright. It is produced through a collaboration
CC       between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC       the European Bioinformatics Institute. There are no restrictions on its
CC       use by non-profit institutions as long as its content is in no way
CC       modified and this statement is not removed. Usage by and for commercial
CC       entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC       or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------
DR   EMBL; M88098; AAA29744.1; -.
KW   Malaria; Receptor; Membrane.
FT   NON_TER       1      1
FT   NON_TER    1251   1251
SQ   SEQUENCE   1251 AA;  143741 MW;  54BA51C7404AC572 CRC64;

Query Match    8.8%;  Score 157.5;  DB 1;  Length 1251;
Best Local Similarity   21.3%;  Pred. No. 2.1;  Mismatches  147;  Indels  157;  Gaps  22;
Matches  102;  Conservative   72;

Qy    9  DGNASANSADESVK-----GPNLTEISKKITESNAVVLAV------KEVELLASIDELAT  58
          ||:| ||| | |        ||||||||||||||        |:|| |||||||
Db  671  EGHGNVKQLVENIKELVDENNLSDILKQATGKNEELQKIIHSTLKNKAKTILGHVDTSA- 729

Qy   59  KAIGKKI----------------GNNGLEA-NQSKNTSLLSGAYAISDLI  91
          | | |                   | | |  |    ||   | |
Db  730  KYVGIKITPELAITELIGDARLKTAQELKFESKNNVVLEFENMSKNTNELDVHKNIQDAY 789

Qy   92  AEKLNVLKNEELKEKIDTAKQCST--------------EFTNKLKS 123
            :  |    | | |  |                      |  |||
Db  790  KVALEILAH--SDEIDTKQKDSSKLIEMGNQIYLKVVLINQYKNKISSIKSKEEAVSVK 846

Qy  124  ------EHAVLG--------LDNLTDDNAQRAILKKHANKDKGAAELE 157
                     |         |  |                  |
Db  847  IGNVSKKHSELSKITCSDKSYDNIIALEKQTELQNLRNSFTQE--KTNTNSD--SKLE 900

Qy  158  KLFKAVENLSKAAQDTLKNAVKELTSPIVHGNNGGKDGNTSANSADESVKGPNLTEISKK 217
          |:| | |  |   |               |        |    ||     |
Db  901  KI------KIDFESLKNALKTLEGE-VNALKASSDNHEHVQSKSEPV-NPALSEIEKE 950

Qy  218  ITESNAVVLAVKEI-----ETLLASIDELATKAIGKKIIQQNGGLAVEAGHNGTLLAGAYTI 273
                    | |         |||||                 |  |              |
Db  951  ETDIDSLNTALDELLKRGRCEVSRYKLLKDTVIKEISDDTEL------------INTI 997

Qy  274  SKLITQKLDGLKNSEKLKEKIENAKKCSEDFTKKLEGEHAQLGIE--NVTDENAKKAILI 331
          ||:          | |:  :| ||          |                  :  |
Db  998  EKNVKAYLAYIK--KNYEDIVQDVLTLNEHFNTKQVSNHEPTNFDKSNKSSEELTKA--V 1053

Qy  332  TDA----AKDKG----------------AAELEKLFKAVENLAKAAKEM--LANSVK 366
          ||       |                    | : |  | : |       |  ::
Db 1054  TDSKTIISKLKGVIIEVNENTEMNTIESSAKEIEALYNELKNKTISLNEIIYQTSNEVK 1111
```

RESULT 14
```
ID   MST2_DROHY     STANDARD;      PRT;   1391 AA.
AC   Q08696;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   AXONEME-ASSOCIATED PROTEIN MST101(2).
GN   MST101(2).
OS   Drosophila hydei (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7224;
RN   [1]
```

RESULT 15
ID   MYSN_DROME   STANDARD;   PRT;   2017 AA.
AC   Q99323;

RP   SEQUENCE FROM N.A., AND CHARACTERIZATION.
RX   MEDLINE=95045538; PubMed=795199;
RA   Neesen J., Padmanabhan S., Buenemann H.;
RT   "Tandemly arranged repeats of a novel highly charged 16-amino-acid
RT   motif representing the major component of the sperm-tail-specific
RT   axoneme-associated protein family Dhmst101 form extended
RT   alpha-helical rods within the extremely elongated spermatozoa of
RT   Drosophila hydel.";
RL   Eur. J. Biochem. 225:1089-1095(1994).
CC   -!- FUNCTION: POSSIBLE STRUCTURAL ROLE IN THE SPERM TAIL.
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC   -!- TISSUE SPECIFICITY: TESTIS. PRIMARY SPERMATOCYTES AND EARLY
CC   SPERMATIDS.
CC   -!- DOMAIN: THE PREDOMINANT STRUCTURE IS ALPHA-HELICAL.
CC   -!- POLYMORPHISM: LENGTH POLYMORPHISMS EXIST BETWEEN DIFFERENT
CC   STRAINS, MOST LIKELY CAUSED BY LENGTH VARIATIONS WITHIN THE TANDEM
CC   REPEATS.

DR   EMBL; X73481; CAA51876.1; -.
DR   PIR; S34154; S34154.
DR   HSSP; P01032; 1C5A.
DR   FlyBase; FBgn0020733; Dhyd\mst101(2).
KW   Sperm; Repeat; Multigene family; Polymorphism.
FT   DOMAIN   332 1268 59 X 16 AA APPROXIMATE TANDEM REPEATS OF
FT   [KR]-K-X-C-X-X-A-K-X-K-X-X-X-E.
SQ   SEQUENCE   1391 AA; 159000 MW; 1B2A36BF30F48878 CRC64;

```
Query Match           8.7%;  Score 157;  DB 1;  Length 1391;
Best Local Similarity 27.3%;  Pred. No. 2.5;
Matches  109;  Conservative  42;  Mismatches  190;  Gaps  17;

Qy    8 KDGNASANSADESVKGPNLTEISKKITESNAVVL-AVKEVETLLASIDELATKAIGKKIG 66
          | |     ||   |||||||||||||| |||    |   ||  |||   | | |
Db  285 KKGKKKKN--DEKKEKELERELLKEQAEEEAKIRGVVKEVK---KKCKEKALKKKCKDLG 339

Qy   67 NNGLEANQSKNTSLLSGAYAISDLIAEKLNYLKNEELKEKIDTAKQCSTEFTNKLKSEHA 126
          |  |   |        :|||:      |:     |: ||| :     ||   ||  ||
Db  340 RKMKEEAEKKKCAALAKKQKEED---EKKACKELAKKKEADEKKKCE-EAANKEKRAAE 395

Qy  127 VLGLDNLTDDNAQRAILKK---HANKRDGAAELEKLFKAVENLSKAAQ-DTLKNAVK--- 179
          |||    |   |||| |||   ||   |   ||   | ||  |   |   ||   || |
Db  396 KKKCEKAAKERKEAAEKKKCEEAAKKEKEAAERKKCEELAKNIKKAAEKKKCKEAAKKEK 455

Qy  180 ------------ELTSPIVHGNNSGKDGNTSANSADESVKGPNLTEISKKITES---NAVVL 226
                     | :|   ||  || |  |  :   |   |||||||||||||  |:   |||
Db  456 EAAERKKCEELAKKIKKAAEKKKCEET-AKKGKEVAERKKCEELAKKIKKAEIKKKCKL 514

Qy  227 AVKEIETLLASIDELATK----AIGKKIQONGGLAVEAGHNGTLLAGAYTISKLITQKL 281
          || |  ||||||||||||    | | |:  ||  | :  || |   :| | | ||  | ||
Db  515 AKKKEKETAEKKKCEKAAAKKRKEAAEKKKCERASAKKREAAEKK- 573

Qy  282 DGLKNSEKLKEKIENA--KKCSEDFTKKLEGEHAQLGIENVTDENAKKAILLTD----- 333
          |   |   |   |:| |  |||    |   |    |   |     |    |   |
Db  574 ---CCEKAAKERKEAAEKKCEE--AAKKEKEVAE---RKKCEELAKKIKKAAEKKKCKE 625

Qy  334 -AAKDKGAAELEKLFKAVENLAKAAK----EMLANSVKE 367
           | :||   | | | |   || | ||    ||| |   |
Db  626 AAKKEKEAAEREKCGELAKKIKKAAEKKKCKKLAKKEKE 664
```

DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   MYOSIN HEAVY CHAIN, NON-MUSCLE (ZIPPER PROTEIN) (MYOSIN II).
GN   ZIP.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RX   MEDLINE=90349606; PubMed=2117279;
RA   Ketchum A.S., Stewart C.T., Stewart M., Kiehart D.P.;
RT   "Complete sequence of the Drosophila nonmuscle myosin heavy-chain
RT   transcript: conserved sequences in the myosin tail and differential
RT   splicing in the 5' untranslated sequence.";
RL   Proc. Natl. Acad. Sci. U.S.A. 87:6316-6320(1990).
CC   -!- FUNCTION: NONMUSCLE MYOSIN APPEARS TO BE RESPONSIBLE FOR
CC   CELLULARIZATION. REQUIRED FOR MORPHOGENESIS AND CYTOKINESIS.
CC   -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
CC   SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 IQ DOMAIN.

DR   EMBL; M35012; AAA28713.1; -.
DR   PIR; A36014; A36014.
DR   PIR; B36014; B36014.
DR   HSSP; P08799; 1MND.
DR   FlyBase; FBgn0005634; zip.
DR   InterPro; IPR000048; IQ.
DR   InterPro; IPR002928; IQ.
DR   InterPro; IPR001609; myosin_head.
DR   InterPro; IPR000612; IQ; 1.
DR   Pfam; PF00063; myosin_head; 1.
DR   Pfam; PF00612; IQ; 1.
DR   Pfam; PF01576; myosin_tail; 1.
DR   PRINTS; PR00193; MYOSINHEAVY.
DR   ProDom; PD000355; myosin_head; 1.
DR   SMART; SM00015; IQ; 1.
DR   SMART; SM00242; MYSc; 1.
DR   PROSITE; PS50096; IQ; 1.
DR   PROSITE; PS50097; 1MD.
KW   Myosin; Alternative splicing; Coiled coil; Actin-binding;
KW   ATP-binding; Calmodulin-binding.
FT   DOMAIN   1   829 MYOSIN HEAD-LIKE.
FT   DOMAIN   830   859 IQ.
FT   NP_BIND   225   232 ATP.
FT   DOMAIN   886 2017 COILED COIL (POTENTIAL).
FT   DOMAIN   250   260 25 KDA/50 KDA JUNCTION.
FT   DOMAIN   682   694 50 KDA/20 KDA JUNCTION.
FT   DOMAIN   705   727 ACTIN-BINDING.
FT   DOMAIN   742   758 REACTIVE SULFHYDRYL/ACTIN-BINDING.
FT   DOMAIN   1303 2017 LIGHT MEROMYOSIN (LMM).
FT   DOMAIN   1303 1970 ALPHA-HELICAL TAILPIECE (LMM).
FT   DOMAIN   1971 2017 GLOBULAR TAILPIECE.
FT   VARSPLIC   1   45 MISSING (IN SHORT ISOFORM).
SQ   SEQUENCE   2017 AA; 232016 MW; 73E3CB02BABF2528 CRC64;

```
Query Match           8.6%;  Score 155.5;  DB 1;  Length 2017;
Best Local Similarity 23.6%;  Pred. No. 4.3;
Matches   91;  Conservative  65;  Mismatches  149;  Gaps  16;

Qy   26 LTEISKKITESNAVV------LAVKEVE---TLLASIDELATKAIGKKIG----------- 66
          :: ::| |  |             |||||   |:|||||||||||||||
Db   91
```

```
Db  1108  VADLKEQLNERRVQVDEMQAQLAKREEELTQTLLRIDEESATKATAQKAQRELESQLAEI  1167
                  :  |   |      |   :::
Qy    67  NNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTAKQCSTEFTNKLKSEHA   126
               |||  |||       :  ::    |   |
Db  1168  QEDLEAEKAARAK----AEKVRRDLSEELEALKNELLDSLDTTPAQ--QELRSKREQELA  1221
                 |         :::|        |||  |||  :   |       |
Qy   127  VLGLDNLFDDNAQRAILKKHAN------KDKGAAELEKLFKRAVENLSKA------AQD   172
               |                   :::  |    :  |||  |
Db  1222  TL--------KKSLEEETVNHEGVLADMRHKHSQELNSINDQLENLRKAKTVLEKAKG  1271
              |  |
Qy   173  TLKNAVKELTSPIVHGNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIE   232
              |  :|:  |     ||| |       :|    :  ||  |      ::
Db  1272  TLEAENADLATELRSVNSSRQENDFRRKQAESQI----AELQVKLAEIERARSELQEKC  1326
              |  :|:  |           |  :           :|
Qy   233  TLLASIDELATKAIGKKIQQNGGLAVEAGHIN---GTLLAGAYTISKLITQKLIDGLKN----   286
              |  ::             |  ::          ||| ||  :
Db  1327  TKLQQEAENITNQL-EEAELKASARAVKSASNMESQLTEAQQLLEEETRQKL-GLSSKLRQ  1384
                 |      |           |  :
Qy   287  ----SEKLKEKIENAKKCSEDFTKKLEGEHAQLGIENVTDENAKKAILITDAAKDKGAAE   342
               :  ||          |  :       |||  |
Db  1385  IESEKEALQEQLEDDEAKRNYERKLAEVTTQM-----QEIKKAEEDADLAK-----E  1433
              |  :|  :       |       |  :       ||||:  :       |
Qy   343  LE-----KLFKRAVENLAKAAKEMLANS   364
              |  :  :    ||||:  ::  | :  :
Db  1434  LEEGKKRLRNKDIEALERQVKELIAQN  1459
```

Search completed: March 18, 2002, 10:11:57
Job time: 975 sec

```
                    GenCore version 4.5
                    Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on:    March 18, 2002, 10:10:55 ; Search time 124.19 Seconds
                                       (without alignments)
                                       433.434 Million cell updates/sec

Title:           US-09-596-746A-42
Perfect score:   1798
Sequence:        1 MACNNSGKDGNASANSADES..........AVENLAKAAKEMLANSVKEL 368

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters:    473505

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :  SPTREMBL_17:*
         1: sp_archea:*
         2: sp_bacteria:*
         3: sp_fungi:*
         4: sp_human:*
         5: sp_invertebrate:*
         6: sp_mammal:*
         7: sp_mhc:*
         8: sp_organelle:*
         9: sp_phage:*
        10: sp_plant:*
        11: sp_rodent:*
        12: sp_virus:*
        13: sp_vertebrate:*
        14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	901	50.1	192	2	Q9RR52	Q9rr52 borrelia bu
2	901	50.1	209	2	Q4717	Q4717 borrelia bu
3	899	50.0	191	2	Q9S3P1	Q9s3p1 borrelia bu
4	889	49.4	192	2	O31117	O31117 borrelia bu
5	876	48.7	193	2	P94234	P94234 borrelia bu
6	876	48.7	194	2	O31122	O31122 borrelia bu
7	876	48.7	201	2	Q9RQR8	Q9rqr8 borrelia bu
8	863	48.0	191	2	O44726	O44726 borrelia bu
9	841	46.8	185	2	O31123	O31123 borrelia bu
10	826	45.9	178	2	Q44995	Q44995 borrelia bu
11	826	45.8	178	2	Q44979	Q44979 borrelia bu
12	823	45.8	178	2	Q44717	Q44717 borrelia bu
13	739	41.1	178	2	Q44994	Q44994 borrelia bu
14	653.5	36.3	212	2	Q44705	Q44705 borrelia bu
15	647.5	36.0	193	2	O31115	O31115 borrelia bu
16	646	35.9	202	2	O08231	O08231 borrelia bu
17	646	35.9	211	2	O44720	O44720 borrelia ta
18	644.5	35.8	192	2	Q9S3P3	Q9s3p3 borrelia bu
19	644.5	35.8	192	2	Q9R7B1	Q9r7b1 borrelia bu
19	644.5	35.8	193	2	Q9RR53	Q9rr53 borrelia ja
20	644.5	35.8	200	2	Q9RQR9	Q9rqr9 borrelia bu
21	642	35.7	193	2	P94237	P94237 borrelia bu
22	641	35.7	188	2	Q9XDH3	Q9xdh3 borrelia bu
23	640.5	35.6	191	2	O31120	O31120 borrelia bu
24	640	35.6	184	2	O31120	O31120 borrelia bu
25	640	35.6	193	2	Q9S504	Q9S504 borrelia bu
26	640	35.6	211	2	P94242	P94242 borrelia bu
27	639.5	35.6	192	2	Q44977	Q44977 borrelia bu
28	639.5	35.6	210	2	Q93P2	Q93p2 borrelia bu
29	637.5	35.5	191	2	Q44719	Q44719 borrelia bu
30	637.5	35.5	201	2	P70818	P70818 borrelia ja
31	635	35.3	201	2	P96572	P96572 borrelia ja
32	634.5	35.3	201	2	Q08138	Q08138 borrelia bu
33	634	35.3	181	2	P96573	P96573 borrelia ja
34	630	35.0	194	2	O34124	O34124 borrelia bu
35	629.5	35.0	201	2	O31114	O31114 borrelia bu
36	629.5	35.0	201	2	P96514	P96514 borrelia ja
37	629	35.0	200	2	P96516	P96516 borrelia ja
38	628.5	35.0	182	2	Q9XDH4	Q9xdh4 borrelia bu
39	628.5	35.0	211	2	Q9R7B2	Q9r7b2 borrelia bu
40	626	34.8	182	2	Q44976	Q44976 borrelia bu
41	625	34.8	184	2	O31121	O31121 borrelia bu
42	625	34.8	193	2	P94236	P94236 borrelia bu
43	624.5	34.7	203	2	O50619	O50619 borrelia bu
44	623.5	34.7	203	2	O50624	O50624 borrelia af
45	623.5	34.7	212	2	Q9KIM5	Q9kim5 borrelia af
	618.5	34.4	201	2	P96571	P96571 borrelia ja

ALIGNMENTS

```
RESULT   1
Q9RR52

ID   Q9RR52                  PRELIMINARY;      PRT;     192 AA.
AC   Q9RR52;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C (FRAGMENT).
OSPC.
GN   OSPC.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=OC5;
RX   MEDLINE=99091544; PubMed=9872945;
RA   Wang I.N., Dykhuizen D.E., Qiu W., Dunn J.J., Bosler E.M., Luft B.J.;
RT   "Genetic diversity of ospC in a local population of Borrelia
RT   burgdorferi sensu stricto.";
RL   Genetics 151:15-30(1999).
DR   EMBL; AF029864; AAB86547.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
FT   NON_TER       1
FT   NON_TER     192
SQ   SEQUENCE   192 AA;   20555 MW;  FF16D409D58C01C5 CRC64;

Query Match            50.1%;   Score 901;   DB 2;   Length 192;
Best Local Similarity  98.9%;   Pred. No. 2.8e-35;
Matches  184;  Conservative  2;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    1 MACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKA 60
          :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7 ISCNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKA 66

Qy   61 IGKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTPAKQCSTEFTNK 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   67 IGKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTPAKQCSTEFTNK 126

Qy  121 LKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKE 180
```

Cross references:
```
Q9rqr9 borrelia bu
P94237 borrelia bu
Q9xdh3 borrelia bu
O31120 borrelia bu
Q9S504 borrelia bu
P94242 borrelia bu
Q44977 borrelia bu
Q93p2 borrelia bu
Q44719 borrelia bu
P70818 borrelia ja
P96572 borrelia ja
Q08138 borrelia bu
P96573 borrelia ja
O34124 borrelia bu
O31114 borrelia bu
P96514 borrelia ja
P96516 borrelia ja
Q9xdh4 borrelia bu
Q9r7b2 borrelia bu
Q44976 borrelia bu
O31121 borrelia bu
P94236 borrelia bu
O50619 borrelia bu
O50624 borrelia af
Q9kim5 borrelia af
P96571 borrelia ja
```

```
Db   127 LKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKE 186

Qy   181 LTSPIV 186
         ||||||
Db   187 LTSPIV 192
```

RESULT 2

```
ID   Q4A717;
AC   Q4A717;
     PRELIMINARY;   PRT;   209 AA.
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
OSPC.
GN   OSPC.
DE   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=N40;
RX   MEDLINE=94314484; PubMed=8039931;
RA   Stevenson B., Bockenstedt L.K., Barthold S.W.;
RT   "Expression and gene sequence of outer surface protein C of Borrelia
RT   burgdorferi reisolated from chronically infected mice.";
RL   Infect. Immun. 62:3568-3571(1994).
DR   EMBL; U04240; AAC45538.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
SQ   SEQUENCE   209 AA;  22393 MW;  3707A47DAA736FCA CRC64;
```

```
Query Match             50.1%;  Score 901;  DB 2;  Length 209;
Best Local Similarity   98.9%;
Matches 184;  Conservative  2;  Mismatches 0;  Indels 0;  Gaps 0;

Qy    1 MACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVAKEVEVETLLASIDELATKA 60
         ::||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   17 ISCNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVAKEVETLLASIDELATKA 76

Qy   61 IGKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTPAKQCSTEFTNK 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   77 IGKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTPAKQCSTEFTNK 136

Qy  121 LKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKE 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  137 LKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKE 196

Qy  181 LTSPIV 186
         ||||||
Db  197 LTSPIV 202
```

RESULT 3

```
ID   Q9S3P1;
AC   Q9S3P1;
     PRELIMINARY;   PRT;   191 AA.
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C (FRAGMENT).
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=28691;
RX   MEDLINE=96256448; PubMed=8709845;
RA   Livey I., Gibbs C.P., Schuster R., Dorner F.;
RT   "Evidence for lateral transfer and recombination in OspC variation in
RT   Lyme disease Borrelia.";
RL   Mol. Microbiol. 18:257-269(1995).
DR   EMBL; L42894; AAB37002.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
FT   NON_TER   1     1
FT   NON_TER   191   191
SQ   SEQUENCE   191 AA;  20340 MW;  8CA1A64CF17AED8F CRC64;
```

```
Query Match             50.0%;  Score 899;  DB 2;  Length 191;
Best Local Similarity   100.0%;
Matches 184;  Conservative  0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy    3 CNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVAKEVEVETLLASIDELATKAIG 62
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 CNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVAKEVEVETLLASIDELATKAIG 60

Qy   63 KKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTPAKQCSTEFTNKLK 122
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 KKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTPAKQCSTEFTNKLK 120

Qy  123 SEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKELT 182
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 SEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKELT 180

Qy  183 SPIV 186
         ||||
Db  181 SPIV 184
```

RESULT 4

```
ID   O31117;
AC   O31117;
     PRELIMINARY;   PRT;   192 AA.
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C (FRAGMENT).
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=OC7;
RA   Wang I.-N., Dykhuizen D.E., Dunn J.J., Luft B.J.;
RL   Submitted (OCT-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF029866; AAB86549.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
FT   NON_TER   1     1
FT   NON_TER   192   192
SQ   SEQUENCE   192 AA;  20684 MW;  1F0C6BB65E291F6B CRC64;
```

```
Query Match             49.4%;  Score 889;  DB 2;  Length 192;
Best Local Similarity   97.8%;
Matches 182;  Conservative  2;  Mismatches 2;  Indels 0;  Gaps 0;

Qy    1 MACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVAKEVEVETLLASIDELATKA 60
         ::||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7 ISCNNSRKDGNASTNSADESVKGPNLTEISKKITESNAVVLAVAKEVETLLASIDELATKA 66

Qy   61 IGKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTPAKQCSTEFTNK 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   67 IGKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKNEELKEKIDTPAKQCSTEFTNK 126

Qy  121 LKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKE 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  127 LKSEHAVLGLDNLTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNAVKE 186
```

```
Qy    181 LTSPIV 186
          |||||
Db    187 LTSPIV 192


RESULT 5
ID      P94234       PRELIMINARY;    PRT;    193 AA.
AC      P94234;
DT      01-MAY-1997 (TrEMBLrel. 03, Created)
DT      01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT      01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE      OUTER SURFACE PROTEIN C (FRAGMENT).
OS      OSPC.
OS      Borrelia burgdorferi (Lyme disease spirochete).
OC      Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX      NCBI_TaxID=139;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=96296448; PubMed=8709845;
RA      Livey I., Gibbs C.P., Schuster R., Dorner F.;
RT      "Evidence for lateral transfer and recombination in OspC variation in
RT      Lyme disease Borrelia.";
RL      Mol. Microbiol. 18:257-269(1995).
DR      EMBL; L42895; AAB37003.1; -.
DR      InterPro; IPR001800; Lipoprotein_6.
DR      Pfam; PF01441; Lipoprotein_6; 1.
DR      ProDom; PD001149; Lipoprotein_6; 1.
NON_TER       1       1
NON_TER     193     193
SEQUENCE    193 AA;  20411 MW;  05B68720F061E2A0 CRC64;


Query Match             48.7%;  Score 876;  DB 2;  Length 193;
Best Local Similarity   100.0%;  Pred. No. 4.1e-34;
Matches  180;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy    189 NNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDELATKAIGK 248
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2 NNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDELATKAIGK  61

Qy    249 KIQQNGGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKEKIENAKKCSEDFTKKL 308
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     62 KIQQNGGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKEKIENAKKCSEDFTKKL 121

Qy    309 EGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLAKAAKEMLANSVKEL 368
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    122 EGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLAKAAKEMLANSVKEL 181


FT      NON_TER       1       1
FT      NON_TER     194     194
SQ      SEQUENCE    194 AA;  20640 MW;  00A5E6D2DCE0F7F CRC64;

RESULT 7
ID      Q9RQR8       PRELIMINARY;    PRT;    201 AA.
AC      Q9RQR8;
DT      01-MAY-2000 (TrEMBLrel. 13, Created)
DT      01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT      01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE      OUTER SURFACE PROTEIN C PRECURSOR (FRAGMENT).
OS      OSPC.
OS      Borrelia burgdorferi (Lyme disease spirochete).
OG      Plasmid cp26.
OC      Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX      NCBI_TaxID=139;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=3B6';
RX      MEDLINE=20002545; PubMed=10531219;
RA      Hofmeister E.K., Glass G.E., Childs J.E., Persing D.H.;
RT      "Population dynamics of a naturally occurring heterogeneous mixture of
RT      Borrelia burgdorferi clones.";
RL      Infect. Immun. 67:5709-5716(1999).
DR      EMBL; AF074465; AAD23912.1; -.
DR      InterPro; IPR001800; Lipoprotein_6.
DR      Pfam; PF01441; Lipoprotein_6; 1.
DR      ProDom; PD001149; Lipoprotein_6; 1.
KW      Plasmid.
FT      NON_TER     201     201
SQ      SEQUENCE    201 AA;  806F19829510 B07 CRC64;


Query Match             48.7%;  Score 876;  DB 2;  Length 201;
Best Local Similarity   100.0%;  Pred. No. 4.3e-34;
Matches  180;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy    189 NNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDELATKAIGK 248
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     20 NNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDELATKAIGK  79

Qy    249 KIQQNGGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKEKIENAKKCSEDFTKKL 308
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     80 KIQQNGGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKEKIENAKKCSEDFTKKL 139

Qy    309 EGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLAKAAKEMLANSVKEL 368
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    140 EGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLAKAAKEMLANSVKEL 199
```

```
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C (FRAGMENT)
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=297;
RX   MEDLINE=94314437; PubMed=8039891;
RA   Fung B.P., McHugh G.L., Leong J.M., Steere A.C.;
RT   "Humoral immune response to outer surface protein C of Borrelia
RT   burgdorferi in Lyme disease: role of the immunoglobulin M response in
RT   the serodiagnosis of early infection.";
RL   Infect. Immun. 62:3213-3221(1994).
DR   EMBL; U08284; AAA21460.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
FT   NON_TER       1      1
SQ   SEQUENCE   191 AA;  20150 MW;  C49A4030F0A28717 CRC64;

Query Match            48.0%;  Score 863;  DB 2;  Length 191;
Best Local Similarity  99.4%;  Pred. No. 1.6e-33;
Matches  178;  Conservative   0;  Mismatches   1;  Indels   0;  Gaps   0;

Qy  190 NSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDELATKAIGKK 249
            |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 NSGKGGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDELATKAIGKK  60

Qy  250 IQQNGGLAVEAGHNGTLLAGAYTISKLITQKLKDGLKNSEKLKEKIENAKKCSEDFTKKLE 309
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 IQQNGGLAVEAGHNGTLLAGAYTISKLITQKLKDGLKNSEKLKEKIENAKKCSEDFTKKLE 120

Qy  310 GEHAQLGIENVTDENAKKAILITDAAKDGKAAELEKLFKAVENLAKAAKEMLANSVKEL 368
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 GEHAQLGIENVTDENAKKAILITDAAKDGKAAELEKLFKAVENLAKAAKEMLANSVKEL 179


RESULT 9
ID   O31123
AC   O31123;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C (FRAGMENT)
OSPC.
GN   OSPC.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=OC13;
RA   Wang I.-N., Dykhuizen D.E., Dunn J.J., Luft B.J.;
RL   Submitted (OCT-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF029872; AAB86555.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6.
DR   ProDom; PD001149; Lipoprotein_6; 1.
FT   NON_TER       1      1
SQ   SEQUENCE   185 AA;  19673 MW;  58D6FEE3C7769CAF CRC64;

Query Match            46.8%;  Score 841;  DB 2;  Length 185;
Best Local Similarity  97.7%;  Pred. No. 1.6e-32;
Matches  172;  Conservative   2;  Mismatches   2;  Indels   0;  Gaps   0;

Qy  189 NNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDELATKAIGK 248
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db   10 NNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASVDELLATKAIGK  69
Qy  249 KIQQNGGLAVEAGHNGTLLAGAYTISKLITQKLKDGLKNSEKLKEKIENAKKCSEDFTKKL 308
            ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db   70 KIQQNGGLAVEAGHNGTLLAGAYTISKLITQELDGLKNSEKLKEKIENAKKCSEDFTKKL 129
Qy  309 EGEHAQLGIENVTDENAKKAILITDAAKDGKAAELEKLFKAVENLAKAAKEMLANS 364
            |||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db  130 EGEHAQLGIENVTDENAKKAILITDAAKDGKAAELEKLFKAVENLAKAAKEFLNNS 185


RESULT 10
ID   Q44995
AC   Q44995;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C (FRAGMENT).
OSPC.
GN   OSPC.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MUL;
RX   MEDLINE=95286481; PubMed=7768799;
RA   Theisen M., Borre M., Mathiesen M.J., Mikkelsen B., Lebech A.M.,
RA   Hansen K.;
RT   "Evolution of the Borrelia burgdorferi outer surface protein OspC.";
RL   J. Bacteriol. 177:3036-3044(1995).
DR   EMBL; X84779; CAA59250.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
FT   NON_TER       1      1
SQ   SEQUENCE   178 AA;  18911 MW;  EEE5OCE48BADFICA CRC64;
```

```
Db   10 NNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASVDELLATKAIGK  69
Qy  249 KIQQNGGLAVEAGHNGTLLAGAYTISKLITQKLKDGLKNSEKLKEKIENAKKCSEDFTKKL 308
Db   70 KIQQNGGLAVEAGHNGTLLAGAYTISKLITQELDGLKNSEKLKEKIENAKKCSEDFTKKL 129
Qy  309 EGEHAQLGIENVTDENAKKAILITDAAKDGKAAELEKLFKAVENLAKAAKEMLANS 364
Db  130 EGEHAQLGIENVTDENAKKAILITDAAKDGKAAELEKLFKAVENLAKAAKEFLNNS 185


RESULT 11
ID   Q44979
AC   Q44979;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C (FRAGMENT).
OSPC.
GN   OSPC.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=272;
RX   MEDLINE=95286481; PubMed=7768799;

Query Match            45.9%;  Score 826;  DB 2;  Length 178;
Best Local Similarity  98.8%;  Pred. No. 7.8e-32;
Matches  171;  Conservative   0;  Mismatches   2;  Indels   0;  Gaps   0;

Qy  196 NTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDELATKAIGKKIQQNG 255
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 NTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDELATKAIGKKIQQNG  60

Qy  256 LAVEAGHNGTLLAGAYTISKLITQKLKDGLKNSEKLKEKIENAKKCSEDFTKKLEGEHAQL 315
            ||||||||||:|||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 LAVEAGHNGTSLAGAYTISKLITQKLKDGLKNSEKLKEKIENAKKCSEDFTKKLEGEHAQL 120

Qy  316 GIENVTDENAKKAILITDAAKDGKAAELEKLFKAVENLAKAAKEMLANSVKEL 368
            |||||||||||||||||||||||||||||||||||||| |||||||||||||
Db  121 GIENVTDENAKKAILITDAAKDGKAAELEKLFKAVENLPKAAKEMLANSVKEL 173
```

RA   Theisen M., Borre M., Mathiesen M.J., Mikkelsen B., Lebech A.M.,
RA   Hansen K.;
RT   "Evolution of the Borrelia burgdorferi outer surface protein OspC.";
RL   J. Bacteriol. 177:3036-3044(1995).
DR   EMBL; X84785; CAA59256.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
FT   NON_TER   1   1
FT   NON_TER   178   178
SQ   SEQUENCE   178 AA;   18938 MW;   972974753EFE81E9 CRC64;

Qy   196   NTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDELATKAIGKKIQQNGG 255
        |||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1   NTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDELATKAIGKKIQQNGG 60

Qy   256   LAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKEKIENAKKCSEDFTKKLEGEHAQL 315
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61   LAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLTEKIENAKRCSEDFTKKLEGEHAQL 120

Qy   316   GIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLAKAAKEMLANSVKEL 368
        ||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db   121   GIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLPKAAKEMLANSVKEL 173

Query Match   45.8%;   Score 823;   DB 2;   Length 178;
Best Local Similarity   98.3%;
Matches 170;   Conservative   1;   Mismatches   2;   Indels   0;   Gaps   0;

RESULT 12
ID   Q44994   PRELIMINARY;   PRT;   178 AA.
AC   Q44994;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C (FRAGMENT).
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95286481; PubMed=7768799;
RA   Theisen M., Borre M., Mathiesen M.J., Mikkelsen B., Lebech A.M.,
RA   Hansen K.;
RT   "Evolution of the Borrelia burgdorferi outer surface protein OspC.";
RL   J. Bacteriol. 177:3036-3044(1995).
DR   EMBL; X84782; CAA59253.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
FT   NON_TER   1   1
FT   NON_TER   178   178
SQ   SEQUENCE   178 AA;   18818 MW;   79B435D0837DA36D CRC64;

Qy   196   NTSANSADESVKGPNLTEISKKITESNAGGLAVKEIETLLASIDELATRAIGKKIQQNGG 255
        |||||||||||||||||||||||||| :::|||||||||||||||||||: |||||||||
Db   1   NTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDELATRAIGKKIQQNGG 60

Qy   256   LAVEAGHNGTLLAGAYTISKLITQKIRWIENSEKITGKIENAKKCSEDFINKLEGEHAQL 315
        ||||||||||||||||||||||||: :|||||:: |||||||||||||||:|||||||||
Db   61   LAVEAGHNGTLSLAGAYTISKLITQKIRWIENSEKITGKIENAKKCSEDFINKLEGEHAQL 120

Qy   316   GIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLAKAAKEMLANSVKEL 368
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121   GIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLAKAAKEMLANSVKEL 368

Query Match   41.1%;   Score 739;   DB 2;   Length 178;
Best Local Similarity   89.0%;   Pred. No. 8.4e-28;
Matches 154;   Conservative   6;   Mismatches   13;   Indels   0;   Gaps   0;

Db   121   GIENVTDENADKAILITDAAKDKGAAELAKLFKAVENLPQAAKEMLANSVKEL 173

RESULT 13
ID   Q44705   PRELIMINARY;   PRT;   212 AA.
AC   Q44705;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94041630; PubMed=8225587;
RA   Padula S.J., Sampieri A., Dias F., Szczepanski A., Ryan R.W.;
RT   "Molecular characterization and expression of p23 (OspC) from a North
RT   American strain of Borrelia burgdorferi.";
RL   Infect. Immun. 61:5097-5105(1993).
DR   EMBL; U01892; AAA16057.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
SQ   SEQUENCE   212 AA;   FB2EF2673A384276 CRC64;

Qy   1   MACNNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATKA 60
        ::|||||||||| ||||||||||||||||||||||||||||||||||||||||:||||||
Db   17   ISCNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDEVAKKA 76

Qy   61   IGKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKN-EELKEKIDTAKQCSTEFTN 119
        | ||| |||| :|| |||:||||| ||||||: ||| |||| ||:||:|
Db   77   IGNLIAQNGLNAGANQNGSLLAGAYVISTLIAERLDGLKNSEELEKIEDAKKCNKAFTD 136

Qy   120   KLKSEHAVLGLDN--LTDDNAQRAILKKHANKDKGAAELEKLFKAVENLSKAAQDTLKNA 177
        ||: | || : ||: | : ::|: |||:||: ||||||||:||||||: : |
Db   137   KLKSSHAELGIANGAASDANAKAAILKTNGTKDKGAQELEKLFESSVKNLSKAAQETLNNS 196

Qy   178   VKELTSPIVHGN 189
        |||||| | |
Db   197   VKELTSPVVAEN 208

Query Match   36.3%;   Score 653.5;   DB 2;   Length 212;
Best Local Similarity   72.4%;   Pred. No. 9.5e-24;
Matches 139;   Conservative   21;   Mismatches   29;   Indels   3;   Gaps   2;

RESULT 14
ID   O31115   PRELIMINARY;   PRT;   193 AA.
AC   O31115;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C (FRAGMENT).
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=OC3;
RA   Wang I.-N., Dykhuizen D.E., Dunn J.J., Luft B.J.;
RL   Submitted (OCT-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF029862; AAB86545.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
FT   NON_TER   1   1

```
FT   NON_TER       193
SQ   SEQUENCE      193 AA;   20596 MW;   018A4CB310475A58 CRC64;

     Query Match             36.0%;  Score 647.5;  DB 2;  Length 193;
     Best Local Similarity   76.7%;
     Matches 138;  Conservative 17;  Mismatches 24;  Indels  1;  Gaps  1;

Qy   189 NNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEIETLLASIDELATKAIGK 248
             |||||||  |||||||||||||||||||||||||||:|||||||||||||||||:|| ||
Db    10 NNSGKDGNASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELA-KAIGK  68

Qy   249 KIQQNGLAVEAGHNGTLLLAGAYTIISKLITQKLDGLKNSEKLKEKIENAKKCSEDFTKKL 308
             ||:| || ||| || ||| ||||||| ||||  || ||||| || | ||||||||:||||
Db    69 KIKNDVSLDNEADNNGSLIAGAYTISTLITQKLSKLNGSEGLKEKIAEAKKCSEEFIKKL 128

Qy   309 EGEHAQLGIENVTDENAKKAILITDAAKDKGAAAELEKLFKAVENLAKAAKEMLANSVKEL 368
             |:|:|||| |:  |:|||||||:|||| | |||||||||||||| ||| ||||| |||||
Db   129 KEKHTDLGKKDATDVHAKEAILKTNGTKDKGAAAELEKLFESVENLAKAAKEMLSNSVKEL 188
```

```
RESULT 15
008231

ID   008231;         PRELIMINARY;        PRT;   202 AA.
AC   008231;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C (FRAGMENT).
OS   Borrelia tanukii.
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=56146;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Fi81T;
RA   Masuzawa T., Komikado T., Fukui T., Yanagihara Y.;
RT   "B. tanukii(Strain Fi81t) ospC, partial cds.";
RL   Submitted (JAN-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB000354; BAA19087.1; -.
DR   InterPro; IPR001800; Lipoprotein.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
FT   NON_TER         1
FT   NON_TER       202
SQ   SEQUENCE      202 AA;  21536 MW;  F309F684C6896lE7 CRC64;

     Query Match             35.9%;  Score 646;  DB 2;  Length 202;
     Best Local Similarity   70.7%;  Pred. No. 2e-23;
     Matches 133;  Conservative 22;  Mismatches 31;  Indels  2;  Gaps  2;

Qy     1 MACNNSGKDG-NASANSADESVKGPNLTEISKKITESNAVVLAVKEVETLLASIDELATK  59
          :||||||||  |||||||||||||||||||||||||||||||||||||||||||||||||
Db    11 ISCNNSGKGGDSASTNPVDESAKGPNLTEISKKITDSNAIVLAVKEVETLISSIDELANK  70

Qy    60 AIGKKIGNNGLEANQSKNTSLLSGAYAISDLIAEKLNVLKN-EELKEKIDTAKQCSTETT 118
          || |||||| |||| ||| || ||||| |:|| ||| ||||  |||||||||||  ||
Db    71 AIGKKIQNNGLDTLSDKNASLLSGAYVISTLITEKLNGLKNSEDLKEKIEEAKKCSAAFT 130

Qy   119 NKLKSEHAVLGLDNLTDDNAQRAILKHANKDKGAAAELEKLFKAVENLSKAAQDTLKNAV 178
          |||:|| |||| ||: |::| ||:|||  |||||||||||||||||| |||| | || |
Db   131 NKLKNSHAQLGIEAATDDDAKKAILKTHVTKDKGAAAEFEKLFKSVEILKAAQEALANSI 190

Qy   179 KELTSPIV 186
          ||||||:|
Db   191 KELTSPVV 198
```

OM protein - protein search, using sw model

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

Run On: March 18, 2002, 09:54:34 ; Search sw model

Sequence: 1 MACNNSGKDGNTSANSADES..........KAAQAALTNSVKELTNPVVA 560
Perfect score: 2750
Title: US-09-596-746A-52

Scoring table: BLOSUM62
Sequence: Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463
Maximum DB seq length: 200000000
Minimum DB seq length: 0

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   A_Geneseq_1101:*
    1: /SIDS8/gcgdata/geneseq/geneseqp/AA1980.DAT:*
    2: /SIDS8/gcgdata/geneseq/geneseqp/AA1981.DAT:*
    3: /SIDS8/gcgdata/geneseq/geneseqp/AA1982.DAT:*
    4: /SIDS8/gcgdata/geneseq/geneseqp/AA1983.DAT:*
    5: /SIDS8/gcgdata/geneseq/geneseqp/AA1984.DAT:*
    6: /SIDS8/gcgdata/geneseq/geneseqp/AA1985.DAT:*
    7: /SIDS8/gcgdata/geneseq/geneseqp/AA1986.DAT:*
    8: /SIDS8/gcgdata/geneseq/geneseqp/AA1987.DAT:*
    9: /SIDS8/gcgdata/geneseq/geneseqp/AA1988.DAT:*
   10: /SIDS8/gcgdata/geneseq/geneseqp/AA1989.DAT:*
   11: /SIDS8/gcgdata/geneseq/geneseqp/AA1990.DAT:*
   12: /SIDS8/gcgdata/geneseq/geneseqp/AA1991.DAT:*
   13: /SIDS8/gcgdata/geneseq/geneseqp/AA1992.DAT:*
   14: /SIDS8/gcgdata/geneseq/geneseqp/AA1993.DAT:*
   15: /SIDS8/gcgdata/geneseq/geneseqp/AA1994.DAT:*
   16: /SIDS8/gcgdata/geneseq/geneseqp/AA1995.DAT:*
   17: /SIDS8/gcgdata/geneseq/geneseqp/AA1996.DAT:*
   18: /SIDS8/gcgdata/geneseq/geneseqp/AA1997.DAT:*
   19: /SIDS8/gcgdata/geneseq/geneseqp/AA1998.DAT:*
   20: /SIDS8/gcgdata/geneseq/geneseqp/AA1999.DAT:*
   21: /SIDS8/gcgdata/geneseq/geneseqp/AA2000.DAT:*
   22: /SIDS8/gcgdata/geneseq/geneseqp/AA2001.DAT:*

Searched: Search time 118.14 Seconds
                 (without alignments)
                 351.118 Million cell updates/sec

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB ID	Description
1	2750	100.0	560	22	AAB62724	Borrelia sp chimer
2	1758	63.9	410	22	AAB62740	Borrelia sp chimer
3	1754	63.8	386	22	AAB62727	Borrelia sp chimer
4	1729.5	62.9	384	22	AAB62726	Borrelia sp chimer
5	1726.5	62.8	408	22	AAB62737	Borrelia sp chimer
6	1491.5	54.2	373	22	AAB62711	Borrelia sp chimer
7	1488.5	54.1	397	22	AAB62729	Borrelia sp chimer
8	1487	54.1	374	22	AAB62710	Borrelia sp chimer
9	1484	54.0	398	22	AAB62728	Borrelia sp chimer
10	1474	53.6	400	22	AAB62739	Borrelia sp chimer
11	1469.5	53.4	401	22	AAB62738	Borrelia sp chimer

ALIGNMENTS

RESULT   1
AAB62724
ID   AAB62724 standard; Protein; 560 AA.
XX
AC   AAB62724;
XX
DT   03-APR-2001 (first entry)
XX
DE   Borrelia sp chimeric ospC protein SEQ ID NO: 52.
XX
KW   Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
XX
PN   WO200078966-A1.
XX
PD   28-DEC-2000.
XX
PF   19-JUN-2000; 2000WO-US16915.
XX
PR   18-JUN-1999;   99US-0140042.
XX
PA   (UYNY) UNIV NEW YORK STATE RES FOUND.
PA   (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI   Dattwyler RJ, Seinost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
DR   WPI; 2001-050113/06.
DR   N-PSDB; AAF29028.
XX
PT   Compositions of OspC polypeptides from strains of Borrelia which cause
PT   Lyme disease are used to immunize animals and detect immune responses
PT   to Lyme disease -

12	1465	53.3	377	22	AAB62713	Borrelia sp chimer
13	1460.5	53.1	378	22	AAB62712	Borrelia sp chimer
14	1452	52.8	378	22	AAB62725	Borrelia sp chimer
15	1444	52.5	401	22	AAB62733	Borrelia sp chimer
16	1198.5	43.6	369	22	AAB62714	Borrelia sp chimer
17	1198.5	43.6	369	22	AAB62716	Borrelia sp chimer
18	1195.5	43.5	393	22	AAB62731	Borrelia sp chimer
19	1195.5	43.5	393	22	AAB62732	Borrelia sp chimer
20	1194	43.4	370	22	AAB62715	Borrelia sp chimer
21	1191	43.3	394	22	AAB62730	Borrelia sp chimer
22	1167.5	42.5	368	22	AAB62717	Borrelia sp chimer
23	1162.5	42.3	367	22	AAB62719	Borrelia sp chimer
24	1159.5	42.2	391	22	AAB62735	Borrelia sp chimer
25	1159.5	42.2	391	22	AAB62736	Borrelia sp chimer
26	1158	42.1	368	22	AAB62718	Borrelia sp chimer
27	1155	42.0	392	22	AAB62734	Borrelia sp chimer
28	986.5	35.9	466	16	AAR75740	B31 outer surface
29	928	33.7	192	22	AAR75741	Borrelia IP2 OspC
30	928	33.7	210	22	AAB62721	B burgdorferi ospC
31	928	33.7	210	16	AAR75727	B. burgdorferi str
32	926	33.7	192	15	AAW11935	Outer surface prot
33	922	33.5	192	22	AAR60886	B31 Osp-A/antigen
34	920	33.5	209	22	AAR75746	Borrelia IP2 OspC
35	915	33.3	466	16	AAR75739	B31 outer surface
36	908	33.0	192	22	AAB62703	B burgdorferi ospC
37	889	32.3	211	22	AAB62722	B burgdorferi ospC
38	889	32.3	212	16	AAR75729	B. garinii ospC pro
39	886	32.2	212	16	AAW41824	B. burgdorferi str
40	885	32.2	206	22	AAB62723	B afzelii ospC pro
41	882	32.1	207	16	AAR75730	B. burgdorferi str
42	880	32.0	194	15	AAR60896	Borrelia JSB antig
43	878	31.9	207	18	AAW41823	B. garinii OspC.
44	864	31.4	176	15	AAR62772	Borrelia B31 antig
45	858	31.2	212	12	AAR13140	B.burgdorferi stra

XX
PS
XX
CC Claim 43; Page 112-113; 160pp; English.
XX
CC The present invention provides compositions comprising ospC proteins and
CC chimeric ospC proteins from members of the Borrelia genus. These may be
CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC vaccines against Borrelia infection, which is spread by ticks and leads
CC to Lyme disease.
XX
SQ Sequence 560 AA;

Query Match Best Local Similarity 100.0%; Score 2750; DB 22; Length 560;
 100.0%; Pred. No. 1.9e-175;
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    1 MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 macnnsgkdgntsansadesvkgpnlteiskkitdsnavllavkeveallssideiaaka 60

Qy   61 IGKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGILKNEGLKEKIDAAKKCSETFTN 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 igkkihqnngldteynhngsllagayaistlikqkldgilknegikekidaakkcsetftn 120

Qy  121 KLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 klkekhtdlgkegvtdadakeailktngtktkgaeelgklfesvevlskaakemlansvk 180

Qy  181 ELTSPVVAESPAMGSNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVET 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 eltspvvaespamgsnsgkggdsastnpadesakgpnlteiskkitdsnafvlavkevet 240

Qy  241 LVLSIDELAKKAIGQKIDNNNGLAALNNQNGSLLAGAYAISTLITEKLSKLKNLEELKTE 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 lvlsidelakkaigqkidnnnglaalnnqngsllagayaistlitekisklknleelkte 300

Qy  301 IAKAKKCSEEFTNKLKSGHADLGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGL 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 iakakkcseeftnklksghadlgkqdatddhakaailkthattdkgakefkdlfesvegl 360

Qy  361 LKAAQVALTNSVKELGHRNNSGGDSDSASTNPDESAKGPNLTVISKITDSNAFLLAVKEVE 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 lkaaqvaltnsvkelghrnnsggdsdsastnpdesakgpnltviskitdsnafllavkeve 420

Qy  421 ALLSSIDELSKAIGKKIKNDGTLDNEANRNESLIAGAYEISKLITQKLSVLNSEELKKKI 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 allssidelskaigkkikndgtldneanrnesliagayeiskiitqkisvlnseelkkki 480

Qy  481 KEAKDCSQKFTTKLKDSHAELGIQSVQDDNAKKAILKTHGTKDKGAKELEELFKSLESLS 540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 keakdcsqkfttklkdshaelgiqsvqddnakkailkthgtkdkgakeleelfkslesls 540

Qy  541 KAAQAALTNSVKELTNPVVA 560
         ||||||||||||||||||||
Db  541 kaaqaaltnsvkeltnpvva 560
```

RESULT 2
ID AAB62740 standard; Protein; 410 AA.
XX
AC AAB62740;
XX
DT 03-APR-2001, (first entry)
XX
DE Borrelia sp chimeric ospC protein SEQ ID NO: 84.
XX
KW Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS Chimeric - Borrelia sp.
OS Chimeric - Borrelia sp.
XX
XX
PN WO200078966-A1.

XX
PD 28-DEC-2000.
XX
PF 19-JUN-2000; 2000WO-US16915.
XX
PR 18-JUN-1999; 99US-0140042.
XX
DR WPI; 2001-050113/06.
DR N-PSDB; AAF29044.
XX
PI Dattwyler RJ, Seinost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
PA (UYNY) UNIV NEW YORK STATE RES FOUND.
PA (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
CC Claim 43; Page 155-156; 160pp; English.
XX
CC The present invention provides compositions comprising ospC proteins and
CC chimeric ospC proteins from members of the Borrelia genus. These may be
CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC vaccines against Borrelia infection, which is spread by ticks and leads
CC to Lyme disease.
XX
SQ Sequence 410 AA;

Query Match Best Local Similarity 95.1%; Score 1758; DB 22; Length 410;
 95.1%; Pred. No. 1.8e-109;
Matches 366; Conservative 0; Mismatches 7; Indels 12; Gaps 1;

```
Qy  188 AESPAMGSNSGKGGDSASTNPDESAKGPNLTEISKKITDSNAFVLAVKEVETLVLSIDE 247
         |||  |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   20 aesigscsnsgkggdsastnpadesakgpnlteiskkitdsnafvlavkevetlvlside 79

Qy  248 LAKKAIGQKIDNNNGLAALNNQNGSLLAGAYAISTLITEKLSKLKNLEELKTEIAKAKKC 307
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   80 lakkaigqkidnnnglaalnnqngsllagayaistlitekisklknleelkteiakakkc 139

Qy  308 SEEFTNKLKSGHADLGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGLLKAAQVA 367
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  140 seeftnklksghadlgkqdatddhakaailkthattdkgakefkdlfesvegllkaaqva 199

Qy  368 LTNSVKELGH--------------RNNSGDSASTNPDESAKGPNLTVISKKITDSNAFLLA 415
         ||||||||                 ||||| || |||||||||||||||||||||||||
Db  200 ltnsvkeltspvvaespkkphmannsggdsastnpdesakgpnltviskkitdsnafila 259

Qy  416 VKEVEALLSSIDELSKAIGKKIKNDGTLDNEANRNESLIAGAYEISKLITQKLSVLNSEE 475
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  260 vkevealissidelskaigkkikndgtldneanrnesliagayeiskiitqkisvlnsee 319

Qy  476 LKKKIKEAKDCSQKFTTKLKDSHAELGIQSVQDDNAKKAILKTHGTKDKGAKELEELFKS 535
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  320 lkkkikeakdcsqkfttklkdshaelgiqsvqddnakkailkthgtkdkgakeleelfks 379

Qy  536 LESLSKAAQAALTNSVKELTNPVVA 560
         ||||||||||||||||||||||||
Db  380 leslskaaqaaltnsvkeltnpvva 404
```

RESULT 3
ID AAB62727 standard; Protein; 386 AA.
XX
AC AAB62727;
XX
DT 03-APR-2001 (first entry)
XX
DE Borrelia sp chimeric ospc protein SEQ ID NO: 58.
XX

Left column:

```
KW   Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
XX
PN   WO200078966-A1.
XX
PD   28-DEC-2000.
XX
PF   19-JUN-2000; 2000WO-US16915.
XX
PR   18-JUN-1999; 99US-0140042.
XX
PA   (UYNY ) UNIV NEW YORK STATE RES FOUND.
PA   (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI   Dattwyler RJ, Seinost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
DR   WPI: 2001-050113/06.
DR   N-PSDB: AAF29031.
XX
PS   Claim 43; Page 120-121; 160pp; English.
XX
CC   Compositions of OspC polypeptides from strains of Borrelia which cause
CC   Lyme disease are used to immunize animals and detect immune responses
CC   to Lyme disease -
XX
CC   The present invention provides compositions comprising ospc proteins and
CC   chimeric ospc proteins from members of the Borrelia genus. These may be
CC   Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC   vaccines against Borrelia infection, which is spread by ticks and leads
CC   to Lyme disease.
XX
SQ   Sequence   386 AA;

Query Match        63.8%;
Best Local Similarity   95.8%;   Score 1754;  DB 22;  Length 386;
Matches 364;  Conservative 0;  Mismatches 4;  Indels 12;  Gaps 1;

Qy  193 MGSNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETLVLSIDELAKKA 252
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 mcsnsgkggdsastnpadesakgpnlteiskkitdsnafvlavkevetlvlsidelakka 60

Qy  253 IGQKIDNNNGLAALNQNGSLLAGAYAISTLITEKLSKLKNLEELKTEIAKAKKCSEEFT 312
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 igqkidnnnglaalnqngsllagayaistlitekls klkneelkteiakakkcseeft 120

Qy  313 NKLKSGHADLGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGLLKAAQVALTNSV 372
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 nklksghadlgkqdatddhakaailkthattdkgakefkdlfesvegllkaaqvaltnsv 180

Qy  373 KELGH-----      RNNSGGDSASTNPDESAKGPNLTVISKKITDSNAFLLAVKEVE 420
            ||                 |||||||||||||||||||||||||||||||||||||||||
Db  181 keltspvvaespkkphmannsggdsastnpdesakgpnltviskkitdsnafllavkeve 240

Qy  421 ALLSSIDELSKAIGKKIKNDGTLDNEANRNESLIAGAYEISKLITQKLSVLNSEELKKKI 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 allssidelskaigkkikndgtldneanrnesliagayeiskiitqkisvlnseelkkki 300

Qy  481 KEAKDCSQKFTTKLKDSHAELGIQSVQDDNAKKAILKTHGTKDKGAKELEELFKSLESLS 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 keakdcsqkfttklkdshaelgiqsvqddnakkailkthgtkdkgakeleelfkslesls 360

Qy  541 KAAQAALTNSVKELTNPVVA 560
            ||||||||||||||||||||
Db  361 kaaqaaltnsvkeltnpvva 380
```

Bottom left:

```
RESULT   4
ID   AAB62726 standard; protein; 384 AA.
XX
```

Right column:

```
AC   AAB62726;
XX
DT   03-APR-2001  (first entry)
XX
DE   Borrelia sp chimeric ospC protein SEQ ID NO: 56.
XX
KW   Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
XX
PN   WO200078966-A1.
XX
PD   28-DEC-2000.
XX
PF   19-JUN-2000; 2000WO-US16915.
XX
PR   18-JUN-1999; 99US-0140042.
XX
PA   (UYNY ) UNIV NEW YORK STATE RES FOUND.
PA   (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI   Dattwyler RJ, Seinost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
DR   WPI: 2001-050113/06.
DR   N-PSDB: AAF29030.
XX
PS   Claim 43; Page 117-118; 160pp; English.
XX
CC   Compositions of OspC polypeptides from strains of Borrelia which cause
CC   Lyme disease are used to immunize animals and detect immune responses
CC   to Lyme disease -
XX
CC   The present invention provides compositions comprising ospc proteins and
CC   chimeric ospc proteins from members of the Borrelia genus. These may be
CC   Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC   vaccines against Borrelia infection, which is spread by ticks and leads
CC   to Lyme disease.
XX
SQ   Sequence   384 AA;

Query Match        62.9%;
Best Local Similarity   66.9%;   Score 1729.5;  DB 22;  Length 384;
Matches 374;  Conservative 1;  Mismatches 3;  Indels 181;  Gaps 1;

Qy    2 ACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAI 61
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 acnnsgkdgntsansadesvkgpnlteiskkitdsnavllavkeveallssideiaakai 60

Qy   62 GKKIHQNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNK 121
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 gkkihqngldteynhngsllagayaistliikqkldglkneglkekidaakkcsetftnk 120

Qy  122 LKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 181
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 lkekhtdlgkegvtdadakeailktngtktkgaeelgklfesvevlskaakemlansvke 180

Qy  182 LTSPVVAESPAMGSNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETL 241
            ||||||||||
Db  181 ltspvvaesp------------------------------------------------ 190

Qy  242 VLSIDELAKKAIGQKIDNNNGLAALNQNGSLLAGAYAISTLITEKLSKLKNLEELKTEI 301
Db  191 ----------------------------------------------------------- 190

Qy  302 AKAKKCSEEFTNKLKSGHADLGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGLL 361
Db  191 ----------------------------------------------------------- 190

Qy  362 KAAQVALTNSVKELGHRNNSGGDSASTNPDESAKGPNLTVISKKITDSNAFLLAVKEVEA 421
            |:   |   |||||||||||||||||||||||||||||||||||||||||
Db  191 ----kkpfhgnnsggdsastnpdesakgpnltviskkitdsnafllavkevea 239
```

RESULT 5

AAB62737
ID   AAB62737 standard; Protein; 408 AA.
XX
AC   AAB62737;
XX
DT   03-APR-2001 (first entry)
XX
DE   Borrelia sp chimeric ospc protein SEQ ID NO: 78.
XX
KW   Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.
XX
OS   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
XX
PN   WO200078966-A1.
XX
PD   28-DEC-2000.
XX
PF   19-JUN-2000; 2000WO-US16915.
XX
PR   18-JUN-1999; 99US-0140042.
XX
PA   (UYNY) UNIV NEW YORK STATE RES FOUND.
PA   (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI   Dattwyler RJ, Seinost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
DR   WPI; 2001-050113/06.
DR   N-PSDB; AAF29041.
XX
PT   Compositions of OspC polypeptides from strains of Borrelia which cause
PT   Lyme disease are used to immunize animals and detect immune responses
PT   to Lyme disease -
XX
PS   Claim 43; Page 147-148; 160pp; English.
XX
SQ   Sequence 408 AA;

Query Match   62.8%; Score 1726.5; DB 22; Length 408;
Best Local Similarity   66.7%; Pred. No. 2.2e-107;
Matches 373; Conservative   2; Mismatches   3; Indels 181; Gaps   1;

Qy   2 ACNNSGKDGNTSANSADESVKGPNCTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAI 61
         :|||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db   25 scnnsgkdgntsansadesvkgpnlteiskkitdsnavllavkeveallssideiaakai 84

Qy   62 GKKIHQNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNK 121
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   85 gkkihqngldteynhngsllagayaistliikqkldglkneglkekidaakkcsetftnk 144

Qy 122 LKEKHTDLGKEGVTDADAKEAILKTNGTRTKGAEELGKLFESVEVLSKAAKEMLANSVKE 181

Qy 422 LLSSIDELSKAIGKKIKNDGTLDNEANRNESLIAGAYEISKLITQKLSVLNSEELKKKIK 481
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 240 lissidelskaigkkikndgtldneanrnesliagayeisklitqklsvlnseelkkkik 299

Qy 482 EAKDCSQKFTTKLKDSHAELGIQSVQDDNAKKAIIKTHGTKDKGAKELEELFKSLESLSK 541
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 300 eakdcsqkfttklkdshaelgigsvqddnakkaiikthgtkdkgakeleelfksleslsk 359

Qy 542 AAQAALTNSVKELTNPVVA 560
         |||||||||||||||||||
Db 360 aaqaaltnsvkeltnpvva 378

RESULT 6

AAB62711
ID   AAB62711 standard; Protein; 373 AA.
XX
AC   AAB62711;
XX
DT   03-APR-2001 (first entry)
XX
DE   Borrelia sp chimeric ospc protein SEQ ID NO: 26.
XX
KW   Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.
XX
OS   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
XX
PN   WO200078966-A1.
XX
PD   28-DEC-2000.
XX
PF   19-JUN-2000; 2000WO-US16915.
XX
PR   18-JUN-1999; 99US-0140042.
XX
PA   (UYNY) UNIV NEW YORK STATE RES FOUND.
PA   (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI   Dattwyler RJ, Seinost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
DR   WPI; 2001-050113/06.
DR   N-PSDB; AAF29015.
XX
PT   Compositions of OspC polypeptides from strains of Borrelia which cause
PT   Lyme disease are used to immunize animals and detect immune responses
PT   to Lyme disease -
XX
PS   Claim 43; Page 81; 160pp; English.

Qy 122 LKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEEIGKLFESVEVLSKAAKEMLANSVKE
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 145 lkekhtdlgkegvtdadakeailktngtktkgaeeigklfesvevlskaakemlansvke 204

Qy 182 LTSPVVAESPAMGSNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETL 241
         ||||||||||
Db 205 ltspvvaesp--- 214

Qy 242 VLSIDELAKKAIGQKIDNNNGLAALNNQNGSLLAGAYAISTLITEKLSKLNEELKTEI 301

Db 215 -- 214

Qy 302 AKAKKCSEEFTNKLKSGHADLGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGLL 361

Db 215 -- 214

Qy 362 RAAQVALTNSVKELGHRNNSGGDSASTNPDESAKGPNLTVISKKITDSNAFLLAVKEVEA 421
          |:| |::   kpfhgnnsgdsastnpdesaakgpnltviskkitdsnafllavkevea
Db 215 ------kpfhgnnsgdsastnpdesaakgpnltviskkitdsnafllavkevea 263

Qy 422 LLSSIDELSKAIGKKIKNDGTLDNEANRNESLIAGAYEISKLITQKLSVLNSEELKKKIK 481
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 264 llssidelskaigkkikndgtldneanrnesliagayeisklitqklsvlnseelkkkik 323

Qy 482 EAKDCSQKFTTKLKDSHAELGIQSVQDDNAKKAIIKTHGTKDKGAKELEELFKSLESLSK 541
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 324 eakdcsqkfttklkdshaelgigsvqddnakkaiikthgtkdkgakeleelfksleslsk 383

Qy 542 AAQAALTNSVKELTNPVVA 560
         |||||||||||||||||||
Db 384 aaqaaltnsvkeltnpvva 402

CC   to Lyme disease.
XX
SQ   Sequence   373 AA;

Query Match   54.2%;   Score 1491.5; DB 22;   Length 373;
Best Local Similarity   82.6%;   Pred. No. 9.3e-92;
Matches 309; Conservative 21; Mismatches 43; Indels 1; Gaps 1;

Qy   2 ACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKVEVEALLSSIDEIAAKAI 61
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 acnnsgkdgntsansadesvkgpnlteiskkitdsnavllavkeveallssideiaakai 60

Qy   62 GKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNK 121
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 gkkihqnngldteynhngsllagayaistlikqkldglkneglkekidaakkcsetftnk 120

Qy   122 LKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 181
          |||||||||||||||||||||||:||||||||||||||||||||||||||||||||||||
Db   121 lkekhtdlgkegvtdadakeailktngtktkgaeelgklfesvevlskaakemlansvke 180

Qy   182 LTSPVVAESPAMGSNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETL 241
          |||||||||||||||||||||||||||||||||||||::||||::||||:||||::|||
Db   181 ltspvvaespamgsnsgkggdsastnpadesakgpnlteiskkitesnavvlavkeietl 239

Qy   242 VLSIDELAKRAIGQKIDNNNGLLAALNNQNGSLLAGAYAISTLITTEKLSKLKNLEELKTEI 301
          |||||||:|||||||||||||||||||||||||||||||||||||||||||||||:|||
Db   240 lasidelatkaigkkiqgnglglaveaghngtllageytiskiltqkldglknseklkeki 299

Qy   302 AKAKKCSEEFTNKLKSGHADLGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGLL 361
          |||||||||||||:||||:||||:|||::||||||||||||:::||:|||||||::|||
Db   300 enakkcsedftkklegehaqlgienvtdenakkalitdaakdkgaaelekifkavenla 359

Qy   362 KAAQVALTNSVKEL 375
          |||||||||:||||
Db   360 kaakemlansvkel 373

RESULT 7
AAB62729
ID   AAB62729 standard; Protein; 397 AA.
XX
AC   AAB62729;
XX
DT   03-APR-2001 (first entry)
XX
DE   Borrelia sp chimeric ospC protein SEQ ID NO: 62.
XX
KW   Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
XX
PN   WO200078966-A1.
XX
PD   28-DEC-2000.
XX
PF   19-JUN-2000; 2000WO-US16915.
XX
PR   18-JUN-1999;   99US-0140042.
XX
PA   (UYNY) UNIV NEW YORK STATE RES FOUND.
PA   (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI   Dattwyler RJ, Seinost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
DR   WPI; 2001-050113/06.
DR   N-PSDB; AAF29033.
XX
PT   Compositions of OspC polypeptides from strains of Borrelia which cause
PT   to Lyme disease -
XX
XX   Compositions of OspC polypeptides from strains of Borrelia which cause
XX   Lyme disease are used to immunize animals and detect immune responses
XX   to Lyme disease.

PS   Claim 43; Page 125-126; 160pp; English.
XX
CC   The present invention provides compositions comprising ospC proteins and
CC   chimeric ospC proteins from members of the Borrelia genus. These may be
CC   chimeric ospC proteins from members of the Borrelia genus. These may be
CC   Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC   vaccines against Borrelia infection, which is spread by ticks and leads
CC   to Lyme disease.
XX
SQ   Sequence   397 AA;

Query Match   54.1%;   Score 1488.5; DB 22;   Length 397;
Best Local Similarity   82.4%;   Pred. No. 1.6e-91;
Matches 308; Conservative 22; Mismatches 43; Indels 1; Gaps 1;

Qy   2 ACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKVEVEALLSSIDEIAAKAI 61
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   25 scnnsgkdgntsansadesvkgpnlteiskkitdsnavllavkeveallssideiaakai 84

Qy   62 GKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNK 121
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   85 gkkihqnngldteynhngsllagayaistlikqkldglkneglkekidaakkcsetftnk 144

Qy   122 LKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 181
          |||||||||||||||||||||||:||||||||||||||||||||||||||||||||||||
Db   145 lkekhtdlgkegvtdadakeailktngtktkgaeelgklfesvevlskaakemlansvke 204

Qy   182 LTSPVVAESPAMGSNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETL 241
          |||||||||||||||||||||||||||||||||||||::||||::||||:||||::|||
Db   205 ltspvvaespamgsnsgkggdsastnpadesakgpnlteiskkitesnavvlavkeietl 263

Qy   242 VLSIDELAKRAIGQKIDNNNGLLAALNNQNGSLLAGAYAISTLITTEKLSKLKNLEELKTEI 301
          |||||||:|||||||||||||||||||||||||||||||||||||||||||||||:|||
Db   264 lasidelatkaigkkiqgnglglaveaghngtllageytiskiltqkldglknseklkeki 323

Qy   302 AKAKKCSEEFTNKLKSGHADLGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGLL 361
          |||||||||||||:||||:||||:|||::||||||||||||:::||:|||||||::|||
Db   324 enakkcsedftkklegehaqlgienvtdenakkalitdaakdkgaaelekifkavenla 383

Qy   362 KAAQVALTNSVKEL 375
          |||||||||:||||
Db   384 kaakemlansvkel 397

RESULT 8
AAB62710
ID   AAB62710 standard; Protein; 374 AA.
XX
AC   AAB62710;
XX
DT   03-APR-2001 (first entry)
XX
DE   Borrelia sp chimeric ospC protein SEQ ID NO: 24.
XX
KW   Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
XX
PN   WO200078966-A1.
XX
PD   28-DEC-2000.
XX
PF   19-JUN-2000; 2000WO-US16915.
XX
PR   18-JUN-1999;   99US-0140042.
XX
PA   (UYNY) UNIV NEW YORK STATE RES FOUND.
PA   (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI   Dattwyler RJ, Seinost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
DR   WPI; 2001-050113/06.

```
DR   N-PSDB; AAF29014.
XX
PT   Compositions of OspC polypeptides from strains of Borrelia which cause
PT   Lyme disease are used to immunize animals and detect immune responses
PT   to Lyme disease -
XX
PS
XX
CC   Claim 43; Page 78-79; 160pp; English.
XX
CC   The present invention provides compositions comprising ospc proteins and
CC   chimeric ospc proteins from members of the Borrelia genus. These may be
CC   Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC   vaccines against Borrelia infection, which is spread by ticks and leads
CC   to Lyme disease.
XX
SQ   Sequence   374 AA;

Query Match          54.1%;   Score 1487;   DB 22;   Length 374;
Best Local Similarity   83.4%;   Pred. No. 1.9e-91;
Matches 312; Conservative 20; Mismatches 40; Indels 2; Gaps 2;

Qy    2 ACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVAKEVEALLSSIDEIAAKAI 61
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 acnnsgkdgntsansadesvkgpnlteiskkitdsnavllavakeveallssideiaakai 60

Qy   62 GKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNK 121
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 gkkihqnngldteynhngsllagayaistlikqkldglkneglkekidaakkcsetftnk 120

Qy  122 LKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 181
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 lkekhtdlgkegvtdadakeailktngtktkgaeelgklfesvevlskaakemlansvke 180

Qy  182 LTSPVVAESPAMGSNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETL 241
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 ltspvvaespamgsnsgkggdsastnpadesakgpnlteiskkitdsnafvlavkevetl 239

Qy  242 VLSIDELAKKAIGQKIDNNNGLAALNNQNGSLLAGAYAISTLITEKLSLKLNLEELKTEI 301
          ||| |||||  ||  :|| :|| :||| ||||||:|| |||||||||| |: :||| ||
Db  240 ltsidela-kaigkkikndvsldneadhngslisgaylisnlitkkisalkdsgelkaei 298

Qy  302 AKAKKCSEEFTNKLKSGHADLGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGLL 361
          :|||||||||| |||:|  ||||:|| ||:||:||||| | |||||| | :|||||| |:
Db  299 ekakkcseeftaklkgehtdlgkegvtddnakkailktnndktkgadeleklfesvknls 358

Qy  362 KAAQVALTNSVKEL 375
          ||| :| ||||||
Db  359 kaakemltnsvkel 372
```

```
RESULT   9
ID   AAB62728 standard; Protein; 398 AA.
XX
AC   AAB62728;
XX
DT   03-APR-2001 (first entry)
XX
DE   Borrelia sp chimeric ospc protein SEQ ID NO: 60.
XX
KW   Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
XX
PN   WO200078966-A1.
XX
PD   28-DEC-2000.
XX
PF   19-JUN-2000; 2000WO-US16915.
XX
PR   18-JUN-1999; 99US-0140042.
XX
```

```
PA   (UYNY ) UNIV NEW YORK STATE RES FOUND.
PA   (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI   Dattwyler RJ, Seinost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
DR   WPI; 2001-050113/06.
DR   N-PSDB; AAF29032.
XX
PT   Compositions of OspC polypeptides from strains of Borrelia which cause
PT   Lyme disease are used to immunize animals and detect immune responses
PT   to Lyme disease -
XX
PS
XX
CC   Claim 43; Page 123; 160pp; English.
XX
CC   The present invention provides compositions comprising ospc proteins and
CC   chimeric ospc proteins from members of the Borrelia genus. These may be
CC   Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC   vaccines against Borrelia infection, which is spread by ticks and leads
CC   to Lyme disease.
XX
SQ   Sequence   398 AA;

Query Match          54.0%;   Score 1484;   DB 22;   Length 398;
Best Local Similarity   83.2%;   Pred. No. 3.2e-91;
Matches 311; Conservative 21; Mismatches 40; Indels 2; Gaps 2;

Qy    2 ACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVAKEVEALLSSIDEIAAKAI 61
          :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   25 scnnsgkdgntsansadesvkgpnlteiskkitdsnavllavakeveallssideiaakai 84

Qy   62 GKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNK 121
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   85 gkkihqnngldteynhngsllagayaistlikqkldglkneglkekidaakkcsetftnk 144

Qy  122 LKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 181
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  145 lkekhtdlgkegvtdadakeailktngtktkgaeelgklfesvevlskaakemlansvke 204

Qy  182 LTSPVVAESPAMGSNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETL 241
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  205 ltspvvaespamgsnsgkggdsastnpadesakgpnlteiskkitdsnafvlavkevetl 263

Qy  242 VLSIDELAKKAIGQKIDNNNGLAALNNQNGSLLAGAYAISTLITEKLSLKLNLEELKTEI 301
          ||| |||||  ||  :|| :|| :||| ||||||:|| |||||||||| |: :||| ||
Db  264 ltsidela-kaigkkikndvsldneadhngslisgaylisnlitkkisalkdsgelkaei 322

Qy  302 AKAKKCSEEFTNKLKSGHADLGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGLL 361
          :|||||||||| |||:|  ||||:|| ||:||:||||| | |||||| | :|||||| |:
Db  323 ekakkcseeftaklkgehtdlgkegvtddnakkailktnndktkgadeleklfesvknls 382

Qy  362 KAAQVALTNSVKEL 375
          ||| :| ||||||
Db  383 kaakemltnsvkel 396
```

```
RESULT  10
ID   AAB62739 standard; Protein; 400 AA.
XX
AC   AAB62739;
XX
DT   03-APR-2001 (first entry)
XX
DE   Borrelia sp chimeric ospc protein SEQ ID NO: 82.
XX
KW   Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
XX
PN   WO200078966-A1.
XX
```

PD   28-DEC-2000.
XX
PF   19-JUN-2000; 2000WO.
XX
PR   18-JUN-1999;   99US-0140042.
XX
PN   (UYNY) UNIV NEW YORK STATE RES FOUND.
PA   (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI   Dattwyler RJ, Seinost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
PR   2001-050113/06.
DR   N-PSDB: AAF29043.
XX
PT   Compositions of OspC polypeptides from strains of Borrelia which cause
PT   Lyme disease are used to immunize animals and detect immune responses
PT   to Lyme disease -
XX
PS   Claim 43; Page 153; 160pp; English.
XX
CC   The present invention provides compositions comprising ospc proteins and
CC   chimeric ospc proteins from members of the Borrelia genus. These may be
CC   vaccines against Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC   to Lyme disease.   Borrelia infection, which is spread by ticks and leads
XX
SQ   Sequence   400 AA;

Query Match   53.6%;   Score 1474;   DB 22;   Length 400;
Best Local Similarity   81.4%;   Pred. No. 1.5e-90;
Matches 307; Conservative 23; Mismatches 43; Indels 4; Gaps 2;

Qy   2 ACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEARLSSSIDEIAAKAI 61
          :||||||||||||||||||||||||||||||||||||||||| |||||: |||:::||
Db   25 scnnsgkdgntsansadesvkgpnlteiskkitdsnavllavkeveallssideiaakai 84

Qy   62 GKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNK 121
          |||||||||||||||||||||||||||||||||||| |||||| ||||||||||||||
Db   85 gkkihqnngldteynhngsllagayaistlikqkldglkneglkekidaakkcsetftnk 144

Qy   122 LKEKHHTDLGKEGVTDADAKEAILKTNGTKTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 181
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   145 lkekhtdlgkegvtdadakeailktngtktkgaeelgklfesvevlskaakemlansvke 204

Qy   182 LTSPVVAES---PAMGSNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEV 238
          :|||||||| :||||||||||||||||||||| ::|| :|||||||||||||||||:
Db   205 ltspvvaespkkpsmvnnsgkdgnt-sansadesvkgpnlteiskkitesnavvlavkei 263

Qy   239 ETLVLSIDELAKKAIGQKIDNNNGLAALNNQNGSLLAGAYAISTLITEKLSKLKNLEELK 298
          ||||||||||||||||||||||||||||||||||||||||||||| ||| ||||||
Db   264 etliasidelatkaigkkiqnnglaveaghngtllagaytisklitcqkidglknsekik 323

Qy   299 TEIAKAKKCSEEFTNKLKSGHADLGKODATDDHAKAAILKTHATTDKGAKEFKDLFESVE 358
          :|| ||||| |||| || |: |||: |||| ||| |||||||||||||||||| |||||
Db   324 ekienakkcseeftkklegehaqlgienvtdenakkailitdaakdkgaaeleklfkave 383

Qy   359 GLLKAAQVALTNSVKEL 375
          || |||: | |||||||
Db   384 nlakaekmlansvkel 400

RESULT 11
ID   AAB62738
AC   AAB62738 standard; Protein; 401 AA.
XX
DT   03-APR-2001   (first entry)
XX
DE   Borrelia sp chimeric ospc protein SEQ ID NO: 80.
XX
KW   Borrelia; ospc; Lyme disease; vaccine; chimeric protein; tick.

XX
OS   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
XX
PN   WO20007896-A1.
XX
PD   28-DEC-2000.
XX
PF   19-JUN-2000; 2000WO-US16915.
XX
PR   18-JUN-1999;   99US-0140042.
XX
PF   (UYNY) UNIV NEW YORK STATE RES FOUND.
PA   (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI   Dattwyler RJ, Seinost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
PR   2001-050113/06.
DR   N-PSDB: AAF29042.
XX
PT   Compositions of OspC polypeptides from strains of Borrelia which cause
PT   Lyme disease are used to immunize animals and detect immune responses
PT   to Lyme disease -
XX
PS   Claim 43; Page 150-151; 160pp; English.
XX
CC   The present invention provides compositions comprising ospc proteins and
CC   chimeric ospc proteins from members of the Borrelia genus. These may be
CC   vaccines against Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC   to Lyme disease.   Borrelia infection, which is spread by ticks and leads
XX
SQ   Sequence   401 AA;

Query Match   53.4%;   Score 1469.5;   DB 22;   Length 401;
Best Local Similarity   82.2%;   Pred. No. 3e-90;
Matches 310; Conservative 22; Mismatches 40; Indels 5; Gaps 3;

Qy   2 ACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEARLSSSIDEIAAKAI 61
          :||||||||||||||||||||||||||||||||||||||||| |||||: |||:::||
Db   25 scnnsgkdgntsansadesvkgpnlteiskkitdsnavllavkeveallssideiaakai 84

Qy   62 GKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNK 121
          |||||||||||||||||||||||||||||||||||| |||||| ||||||||||||||
Db   85 gkkihqnngldteynhngsllagayaistlikqkldglkneglkekidaakkcsetftnk 144

Qy   122 LKEKHHTDLGKEGVTDADAKEAILKTNGTKTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 181
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   145 lkekhtdlgkegvtdadakeailktngtktkgaeelgklfesvevlskaakemlansvke 204

Qy   182 LTSPVVAES---PAMGSNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEV 238
          :|||||||| :||||||||||||||||||||| ::|| :|||||||||||||||||:
Db   205 ltspvvaespkkpsmvnnsgkdgnt-sansadesvkgpnlteiskkitesnavvlavkev 263

Qy   239 ETLVLSIDELAKKAIGQKIDNNNGLAALNNQNGSLLAGAYAISTLITEKLSKLKNLEELK 298
          ||||||||||||||||||||||||||||||||||||||||||||| ||| ||||||
Db   264 etiltsidela-kaigkkndvsldneadhngsllsgaylisnlitkkisalkdsgelk 322

Qy   299 TEIAKAKKCSEEFTNKLKSGHADLGKODATDDHAKAAILKTHATTDKGAKEFKDLFESVE 358
          |||| ||||||| |||| ||| ||| |||:::|:|:||||||||||||||||||:
Db   323 aeiekakkcseeftaklkgehtdlgkegvtddnakkailktnndktkgadeleklfesvk 382

Qy   359 GLLKAAQVALTNSVKEL 375
          || |||||| ||||||
Db   383 nlskaaekmltnsvkel 399

RESULT 12
ID   AAB62713
AC   AAB62713 standard; Protein; 377 AA.
XX
AC   AAB62713;

XX
DT   03-APR-2001 (first entry)
XX
DE   Borrelia sp chimeric ospC protein SEQ ID NO: 30.
XX
KW   Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
XX
PN   WO200078966-A1.
XX
PD   28-DEC-2000.
XX
PF   19-JUN-2000; 2000WO-US16915.
XX
PR   18-JUN-1999; 99US-0140042.
XX
PA   (UYNY) UNIV NEW YORK STATE RES FOUND.
PA   (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI   Dattwyler RJ, Seinost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
DR   WPI; 2001-050113/06.
DR   N-PSDB; AAF29017.
XX
PT   Compositions of OspC polypeptides from strains of Borrelia which cause
PT   Lyme disease are used to immunize animals and detect immune responses
PT   to Lyme disease -
XX
PS   Claim 43; Page 86-87; 160pp; English.
XX
CC   The present invention provides compositions comprising ospC proteins and
CC   chimeric ospC proteins from members of the Borrelia genus. These may be
CC   Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC   vaccines against Borrelia infection, which is spread by ticks and leads
CC   to Lyme disease.
XX
SQ   Sequence   377 AA;

Query Match   53.3%; Score 1465; DB 22; Length 377;
Best Local Similarity   81.0%; Pred. No. 5.5e-90;
Matches 306; Conservative   23; Mismatches   45; Indels   4; Gaps   2;

Qy   1 MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 macnnsgkdgntsansadesvkgpnlteiskkitdsnavllavkeveallssideiaaka 60

Qy   61 IGKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN 120
        |||||||||||||:|||:|||||||||||||||||||||||||||||||||||||||||
Db   61 igkkihqnngldtennhngsllagayaistllikqkldglkneglkekidaakkcsetftn 120

Qy 121 KLKEKHTDLGKEGVTDADAKEAILKTNGTRTKGAEELGKLFESVEVLSKAAKEMLANSVK 180
        ||||||||||||||||||||||||||:||:||||||||||||||||||||||||||||||
Db 121 klkekhtdlgkegvtdadakeailkangtktkgaeelgklfesvevlskaakemlansvk 180

Qy 181 ELTSPVVAES---PAMGSNSGKGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKE 237
        ||||||||| : |:|| ::|| || |||||||||:||||||||||:||:|| :|||||
Db 181 eltspvvaespkkpsmvnnsgkdgnt-sansadesvkgpnlteiskkitesnavvlavke 239

Qy 238 VETLVLSIDELAKKAIGQKIDNNNGLAALNNQNGSLLAGAYAISTLITEKLSKLKNLEEL 297
        ||| : |:|||| :||||:| : :|||||: ||||||||||||||||:|| |:|:|||
Db 240 ietllasidelatkaigkkiqqnnggllagayaistlikqt1iqkldglknsekl 299

Qy 298 KTEIAKAKKCSEEFTNKLKSGHADLGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESV 357
        |||| |||||| ||||:|:|| || ||||||||:|||||||:||:|||||||||||:||
Db 300 kekienakkcsedftkklegehaqlgienvtdenakkalitdaakdkgaaelek1fkav 359

Qy 358 EGLLKAAQVALTNSVKEL 375
        |||:|| :| ||||||:|
Db 360 eglakkaaqvaltnsvkel 377

Db   299 kaeiekakkcseeftaklkgehtdlgkegvtddnakkailktnndktkgadelekifesv 358
Qy   358 EGLLKAAQVALTNSVKEL 375
         :||||: |||||||||
Db   359 knlskaakemltnsvkel 376

RESULT 14
AAB62725
ID   AAB62725 standard; Protein; 378 AA.
XX
AC   AAB62725;
XX
DT   03-APR-2001 (first entry)
XX
DE   Borrelia sp chimeric ospc protein SEQ ID NO: 54.
XX
KW   Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
XX
PN   WO200078966-A1.
XX
PD   28-DEC-2000.
XX
PR   18-JUN-1999; 99US-0140042.
XX
PF   19-JUN-2000; 2000WO-US16915.
XX
PA   (UYNY) UNIV NEW YORK STATE RES FOUND.
PA   (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI   Dattwyler RJ, Selnost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
PT   Lyme disease -
PT   Lyme disease are used to immunize animals and detect immune responses
PN
XX
CC   Claim 43; Page 114-115; 160pp; English.
XX
PS
PD
XX
CC   Compositions of OspC polypeptides from strains of Borrelia which cause
CC   Lyme disease are used to immunize animals and detect immune responses
CC   to Lyme disease.
XX
CC   The present invention provides compositions comprising ospc proteins and
CC   chimeric ospc proteins from members of the Borrelia genus. These may be
CC   Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC   vaccines against Borrelia infection, which is spread by ticks and leads
CC   to Lyme disease.
XX
DR   WPI; 2001-050113/06.
DR   N-PSDB; AAF29029.
XX

Sequence   378 AA;

Query Match   52.8%; Score 1452; DB 22; Length 378;
Best Local Similarity 58.7%; Pred. No. 4.1e-89;
Matches 330; Conservative   9; Mismatches 31; Indels 192; Gaps   4;

Qy   1 MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA 60
         ||||| : |||:|||||:|||||||||:||||||||||||||||||||||:|| |||
Db   1 macnnsgkdgntsansadesvkgppnlteiskkitdsnavllavkeveallssidei-aka 59

Qy   61 IGKKIHQNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLK-NEGLKEKIDAARKCSETFT 119
         ||||| |: |||||||||||||||:| :||||||: :||||||||||:
Db   60 igkkikndgsldneanrnesllagaytistlitqkislngseglkekiaaakkcseefs 119

Qy   120 NKLKEKHTDLGKEGVTDADAKEAILKTNGT-KTKGAEELGKLFESVEVLSKAAKEMLANS 178
         :||: |||| ::|||| :|:|||| |:||| |||||||||||||||||||||||||||
Db   120 tklkdnhaqlgiqgvtdenakkailkanaagkdkgveelekisgslesiskaakemlans 179

Qy   179 VKELTSPVVAESPAMGSNSNGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEV 238
         ||||||||
Db   180 vkeltspvv--- 188

Qy   239 ETLVLSIDELAKKAIGQKIDNNGLAALNNQNGSLLAGAYAISTLIITEKLSKLRNLEELK 298
Db   189 --- 188

Qy   299 TEIAKAKKCSEEFTNKLKGSHADLGKQDATDDHAKAAILKTHAFTDKGAKEFKDLFPESVE 358
Db   189 -- 188

Qy   359 GLLKAAQVALTNSVKELGHRNNSGGDSASTNPDESAKGPNLTVISKKITDSNAFLLAVKE 418
           |||||||||||||| ||||||||||||||||||||||||||||||||||
Db   189 -----------hgnnsggdsastnpdesakgpnltviskkitdsnafllavke 230

Qy   419 VEALLSSIDELSKRAIGKKIKNDGTLDNEANRNESLIAGAYEISKLIITOKLSVLNSEELKK 478
         ||||||||||||:|||||||||||||||||||||:|||||||||||:||||||||||||||
Db   231 veallssidelskaigkkikndgtldneanrnesliagayeisklitqklsvlnseelkk 290

Qy   479 KIKEAKDCSQKFTPKLKDSHAELGIQSVQDDNAKKAILKTHGTKDKGAKELEELFKSLES 538
         |||||||||:|||||||||||||||||||||||||||||||||||||||||||||||
Db   291 kikeakdcsqkfttklkdshaelgiqsvqddnakkailkthgtkdkgakeleelfksles 350

Qy   539 LSKAAQOAALTNSVKELTNPVVA 560
         |||||||||||||||||||||
Db   351 lskaaqaaltnsvkeltnpvva 372

RESULT 15
AAB62733
ID   AAB62733 standard; Protein; 401 AA.
XX
AC   AAB62733;
XX
DT   03-APR-2001 (first entry)
XX
DE   Borrelia sp chimeric ospc protein SEQ ID NO: 70.
XX
KW   Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS   Chimeric - Borrelia sp.
OS   Chimeric - Borrelia sp.
XX
PN   WO200078966-A1.
XX
PD   28-DEC-2000.
XX
PR   18-JUN-1999; 99US-0140042.
XX
PF   19-JUN-2000; 2000WO-US16915.
XX
PA   (UYNY) UNIV NEW YORK STATE RES FOUND.
PA   (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI   Dattwyler RJ, Selnost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
PT   Lyme disease -
PT   Lyme disease are used to immunize animals and detect immune responses
PN
XX
CC   Claim 43; Page 136-137; 160pp; English.
XX
PS
XX
CC   Compositions of OspC polypeptides from strains of Borrelia which cause
CC   Lyme disease are used to immunize animals and detect immune responses
CC   to Lyme disease.
XX
CC   The present invention provides compositions comprising ospc proteins and
CC   chimeric ospc proteins from members of the Borrelia genus. These may be
CC   Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC   vaccines against Borrelia infection, which is spread by ticks and leads
CC   to Lyme disease.
XX
DR   WPI; 2001-050113/06.
DR   N-PSDB; AAF29037.
XX
SQ   Sequence   401 AA;

Query Match   52.5%; Score 1444; DB 22; Length 401;
Best Local Similarity 58.5%; Pred. No. 1.5e-88;

Matches   328;   Conservative   10;   Mismatches   31;   Indels   192;   Gaps   4;

```
Qy    2  ACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAI    61
         :|||||||||||||||||||||||||||||||||||||||||||||: ||||||: | ||||
Db   25  scnnsgkdgntsansadesvkgpnlteiskkitdsnavllavkeveallssidel-akai   83

Qy   62  GKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLK-NEGLKEKIDAAKKCSETFTN   120
         |||:| | ||| |||||||||||| |||||: ||| |||   :|||||||| ||||||:
Db   84  gkkihndgsldneanrnesllagaytistliltqklsklngseglkekiaaakkcseefst  143

Qy  121  KLKEKHTDLGKEGVTDADAKEAILKTNGT-KTKGAEELGKLFESVEVLSKAAKEMLANSV   179
         |||:| :| | | |||| ::|||||:|  ||| |||||| |:|: |||||||||||||
Db  144  klkdnhaqlgiqgyvtdenakkailkanaagkdkgveelekisgsleslskaakemlansv  203

Qy  180  KELTSPVVAESPAMGSNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVE   239
         |||||||||
Db  204  keltspvv----------------------------------------------------  211

Qy  240  TLVLSIDELAKKAIGQKIDNNNGLAALNNQNGSLLAGAYAISTLITEKLSKLKNLEELKT   299
Db  212  -----------------------------------------------------------  211

Qy  300  EIAKAKKCSEEFTNKLKSGHADLGKRQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEG   359
Db  212  -----------------------------------------------------------  211

Qy  360  LLKAAQVALTNSVKELGHRNNSGGDSASTNPDESAKGPNLTVISKKITDSNAFLLAVKEV   419
                   |  |||||||||||||||||||||||||||||||||||||
Db  212  ----------hgnnsggdsastnpdesakgpnltviskkitdsnafllavkev  254

Qy  420  EALLSSIDELSKAIGKKIKNDGTLDNEANRNESLIAGAYEISKLITQKLSVLNSEELKKK   479
         ||||||||||||||||||||||||||||||||||||||:|||:||||||||||||||||
Db  255  eallssidelskaigkkikndgtldneanrnesllagayeiskiitqklsvlnseelkkk  314

Qy  480  IKEAKDCSQKFTTKLKDSHAELGIOSVQDDNAKKAILKTHGTKDKGAKELEELFKSLESL   539
         ||||||||||||||||||||||||:||||||||||||||||||||||||||||||||:||
Db  315  ikeakdcsqkfttklkdshaelgiqsvqddnakkailkthgtkdkgakeleelfksles1  374

Qy  540  SKAAQAALTNSVKELTNPVVA   560
         |||||||||||||||||||||
Db  375  skaaqaaltnsvkeltnpvva   395
```

OM protein - protein search, using sw model

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

Run on: March 18, 2002, 10:08:46 ; Search time 621.2 Seconds
                                  (without alignments)
                                    250.303 Million cell updates/sec

Title: US-09-596-746A-52
Perfect score: 2750
Sequence: 1 MACNNSGKDGNTSANSADES..........KAAQAALTNSVKELTNPVVA 560

Scoring table: BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched: 3148936 seqs, 277657034 residues

Post-processing: Minimum Match 0%
Maximum DB seq length: 0
Maximum DB seq length: 2000000000

Total number of hits satisfying chosen parameters: 3148936

Database :   Pending_Patents_AA_Main:*
           Minimum Match 100%
           Listing first 45 summaries

   1: /cgn2_6/ptodata/2/paa/PCTUS_COMB.pep:*
   2: /cgn2_6/ptodata/2/paa/US06_COMB.pep:*
   3: /cgn2_6/ptodata/2/paa/US07_COMB.pep:*
   4: /cgn2_6/ptodata/2/paa/US080_COMB.pep:*
   5: /cgn2_6/ptodata/2/paa/US081_COMB.pep:*
   6: /cgn2_6/ptodata/2/paa/US082_COMB.pep:*
   7: /cgn2_6/ptodata/2/paa/US083_COMB.pep:*
   8: /cgn2_6/ptodata/2/paa/US084_COMB.pep:*
   9: /cgn2_6/ptodata/2/paa/US085_COMB.pep:*
  10: /cgn2_6/ptodata/2/paa/US086_COMB.pep:*
  11: /cgn2_6/ptodata/2/paa/US087_COMB.pep:*
  12: /cgn2_6/ptodata/2/paa/US088_COMB.pep:*
  13: /cgn2_6/ptodata/2/paa/US089_COMB.pep:*
  14: /cgn2_6/ptodata/2/paa/US090_COMB.pep:*
  15: /cgn2_6/ptodata/2/paa/US091_COMB.pep:*
  16: /cgn2_6/ptodata/2/paa/US092_COMB.pep:*
  17: /cgn2_6/ptodata/2/paa/US093_COMB.pep:*
  18: /cgn2_6/ptodata/2/paa/US094_COMB.pep:*
  19: /cgn2_6/ptodata/2/paa/US095_COMB.pep:*
  20: /cgn2_6/ptodata/2/paa/US096_COMB.pep:*
  21: /cgn2_6/ptodata/2/paa/US097_COMB.pep:*
  22: /cgn2_6/ptodata/2/paa/US098_COMB.pep:*
  23: /cgn2_6/ptodata/2/paa/US099_COMB.pep:*
  24: /cgn2_6/ptodata/2/paa/US60_COMB.pep:*

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2750	100.0	560	19	US-09-596-746-52	Sequence 52, Appl
2	2750	100.0	560	19	US-09-596-746A-52	Sequence 52, Appl
3	1758	63.9	410	19	US-09-596-746A-84	Sequence 84, Appl
4	1758	63.9	411	19	US-09-596-746A-84	Sequence 84, Appl
5	1754	63.8	386	19	US-09-596-746-58	Sequence 58, Appl
6	1754	63.8	386	19	US-09-596-746A-58	Sequence 58, Appl
7	1734.5	63.1	385	19	US-09-596-746A-56	Sequence 56, Appl
8	1729.5	62.9	384	19	US-09-596-746-56	Sequence 56, Appl
9	1726.5	62.8	408	19	US-09-596-746-78	Sequence 78, Appl

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

ALIGNMENTS

10	1726.5	62.8	409	19	US-09-596-746A-78	Sequence 78, Appl
11	1496.5	54.4	374	19	US-09-596-746A-26	Sequence 26, Appl
12	1492	54.3	375	19	US-09-596-746A-24	Sequence 24, Appl
13	1491.5	54.2	373	19	US-09-596-746-26	Sequence 26, Appl
14	1488.5	54.1	397	19	US-09-596-746-62	Sequence 62, Appl
15	1488.5	54.1	398	19	US-09-596-746A-62	Sequence 62, Appl
16	1487	54.1	374	19	US-09-596-746-24	Sequence 24, Appl
17	1484	54.0	399	19	US-09-596-746-60	Sequence 60, Appl
18	1484	53.6	400	19	US-09-596-746A-60	Sequence 60, Appl
19	1474	53.6	401	19	US-09-596-746A-82	Sequence 82, Appl
20	1469.5	53.4	401	19	US-09-596-746-82	Sequence 82, Appl
21	1469.5	53.4	402	19	US-09-596-746A-80	Sequence 80, Appl
22	1465	53.3	402	19	US-09-596-746-80	Sequence 80, Appl
23	1465	53.3	377	19	US-09-596-746-30	Sequence 30, Appl
24	1460.5	53.1	377	19	US-09-596-746A-30	Sequence 30, Appl
25	1460.5	53.1	378	19	US-09-596-746-28	Sequence 28, Appl
26	1452	52.8	378	19	US-09-596-746A-28	Sequence 28, Appl
27	1452	52.8	378	19	US-09-596-746-54	Sequence 54, Appl
28	1444	52.5	401	19	US-09-596-746-70	Sequence 70, Appl
29	1444	52.5	369	19	US-09-596-746A-54	Sequence 54, Appl
30	1444	52.5	402	19	US-09-596-746A-70	Sequence 70, Appl
31	1203.5	43.8	370	19	US-09-596-746A-32	Sequence 32, Appl
32	1203.5	43.8	370	19	US-09-596-746-32	Sequence 32, Appl
33	1199	43.6	371	19	US-09-596-746A-34	Sequence 34, Appl
34	1198.5	43.6	369	19	US-09-596-746-34	Sequence 34, Appl
35	1198.5	43.6	369	19	US-09-596-746-32	Sequence 32, Appl
36	1195.5	43.5	393	19	US-09-596-746-66	Sequence 66, Appl
37	1195.5	43.5	393	19	US-09-596-746A-68	Sequence 68, Appl
38	1195.5	43.5	394	19	US-09-596-746-66	Sequence 66, Appl
39	1195.5	43.5	394	19	US-09-596-746A-66	Sequence 66, Appl
40	1194	43.4	370	19	US-09-596-746-34	Sequence 34, Appl
41	1191	43.3	394	19	US-09-596-746-64	Sequence 64, Appl
42	1191	43.3	395	19	US-09-596-746A-64	Sequence 64, Appl
43	1167.5	42.5	368	19	US-09-596-746-38	Sequence 38, Appl
44	1167.5	42.5	368	19	US-09-596-746A-38	Sequence 38, Appl
45	1167.5	42.5	368	19	US-09-596-746A-42	Sequence 42, Appl

RESULT 1
US-09-596-746-52
Sequence 52, Application US/09596746
GENERAL INFORMATION:
;   APPLICANT: Dattwyler, Raymond J.
;   APPLICANT: Seinost, Gerald
;   APPLICANT: Dykhuizen, Daniel
;   APPLICANT: Luft, Benjamin J.
;   APPLICANT: Maria J.C. Gomes-Solecki
;   TITLE OF INVENTION: Groups of Borrelia burgdorferi and
;   TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
;   FILE REFERENCE: 2631.1002-001
;   CURRENT APPLICATION NUMBER: US/09/596,746
;   CURRENT FILING DATE: 2000-06-16
;   PRIOR APPLICATION NUMBER: US 60/140,042
;   PRIOR FILING DATE: 1999-06-18
;   NUMBER OF SEQ ID NOS: 84
;   SOFTWARE: FastSEQ for Windows Version 4.0
;   SEQ ID NO 52
;   LENGTH: 560
;   TYPE: PRT
;   ORGANISM: ospC Chimera
US-09-596-746-52

Query Match   100.0%;   Score 2750;   DB 19;   Length 560;
Best Local Similarity   100.0%;   Pred. No. 1.1e-180;
Matches   560;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   1 MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA 60
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA 60

```
RESULT 2
US-09-596-746A-52
; Sequence 52, Application US/0959674A
; GENERAL INFORMATION:
;  APPLICANT: Dattwyler, Raymond J.
;  APPLICANT: Seinost, Gerald
;  APPLICANT: Dykhuizen, Daniel
;  APPLICANT: Luft, Benjamin J.
;  APPLICANT: Maria J.C. Gomes-Solecki
;  TITLE OF INVENTION: Groups of Borrelia burgdorferi and
;  TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
;  FILE REFERENCE: 2631.1002-001
;  CURRENT APPLICATION NUMBER: US/09/596,746A
;  CURRENT FILING DATE: 2000-06-19
;  PRIOR APPLICATION NUMBER: US 60/140,042
;  PRIOR FILING DATE: 1999-06-18
;  NUMBER OF SEQ ID NOS: 84
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 52
;  LENGTH: 560
;  TYPE: PRT
;  ORGANISM: ospC Chimera
US-09-596-746A-52

  Query Match             100.0%;  Score 2750;  DB 19;  Length 560;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-180;
  Matches  560;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    1 MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAKA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAKA 60

Qy   61 IGKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSEFTN 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 IGKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSEFTN 120

Qy  121 KLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 KLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 180

Qy  181 ELTSPVVAESPAMGSNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVET 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 ELTSPVVAESPAMGSNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVET 240

Qy  241 LVLSIDELAKKAIGQKIDNNNGLAALNNQNGSLLAGAYAISTLITEKLSKLKNLEELKTE 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 LVLSIDELAKKAIGQKIDNNNGLAALNNQNGSLLAGAYAISTLITEKLSKLKNLEELKTE 300

Qy  301 IAKAKKCSEEFTNKLKSGHADLGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGL 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 IAKAKKCSEEFTNKLKSGHADLGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGL 360

Qy  361 LKAAQVALTNSVKELGHRNNSGGDSASTNPDESAKGPNLTVISKKITDSNAFLLAVKEVE 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 LKAAQVALTNSVKELGHRNNSGGDSASTNPDESAKGPNLTVISKKITDSNAFLLAVKEVE 420

Qy  421 ALLSSIDELSKAIGKKIKNDGTLDNEANRNESLIAGAYEISKLITQKLSVLNSEELKKKI 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 ALLSSIDELSKAIGKKIKNDGTLDNEANRNESLIAGAYEISKLITQKLSVLNSEELKKKI 480

Qy  481 KEAKDCSQKFTTKLKDSHAELGIQSVQDDNAKKAILKTHGTKDKGAKELEELFKSLESLS 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 KEAKDCSQKFTTKLKDSHAELGIQSVQDDNAKKAILKTHGTKDKGAKELEELFKSLESLS 540

Qy  541 KAAQAALTNSVKELTNPVVA 560
          ||||||||||||||||||||
Db  541 KAAQAALTNSVKELTNPVVA 560
```

```
RESULT 3
US-09-596-746A-84
; Sequence 84, Application US/0959674A
; GENERAL INFORMATION:
;  APPLICANT: Dattwyler, Raymond J.
;  APPLICANT: Seinost, Gerald
;  APPLICANT: Dykhuizen, Daniel
;  APPLICANT: Luft, Benjamin J.
;  APPLICANT: Maria J.C. Gomes-Solecki
;  TITLE OF INVENTION: Groups of Borrelia burgdorferi and
;  TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
;  FILE REFERENCE: 2631.1002-001
;  CURRENT APPLICATION NUMBER: US/09/596,746
;  CURRENT FILING DATE: 2000-06-16
;  PRIOR APPLICATION NUMBER: US 60/140,042
;  PRIOR FILING DATE: 1999-06-18
;  NUMBER OF SEQ ID NOS: 84
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 84
;  LENGTH: 410
;  TYPE: PRT
;  ORGANISM: ospC Chimera
US-09-596-746A-84

  Query Match              63.9%;  Score 1758;  DB 19;  Length 410;
  Best Local Similarity    95.1%;  Pred. No. 1.6e-112;
  Matches  366;  Conservative    0;  Mismatches    7;  Indels   12;  Gaps    1;

Qy   61 IGKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSEFTN 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 IGKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSEFTN 60

Qy  121 KLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 KLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 120

Qy  181 ELTSPVVAESPAMGSNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVET 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 ELTSPVVAESPAMGSNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVET 180

Qy  241 LVLSIDELAKKAIGQKIDNNNGLAALNNQNGSLLAGAYAISTLITEKLSKLKNLEELKTE 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 LVLSIDELAKKAIGQKIDNNNGLAALNNQNGSLLAGAYAISTLITEKLSKLKNLEELKTE 240

Qy  301 IAKAKKCSEEFTNKLKSGHADLGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGL 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 IAKAKKCSEEFTNKLKSGHADLGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGL 300

Qy  361 LKAAQVALTNSVKELGHRNNSGGDSASTNPDESAKGPNLTVISKKITDSNAFLLAVKEVE 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 LKAAQVALTNSVKELGHRNNSGGDSASTNPDESAKGPNLTVISKKITDSNAFLLAVKEVE 360

Qy  421 ALLSSIDELSKAIGKKIKNDGTLDNEANRNESLIAGAYEISKLITQKLSVLNSEELKKKI 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 ALLSSIDELSKAIGKKIKNDGTLDNEANRNESLIAGAYEISKLITQKLSVLNSEELKKKI 420

Qy  481 KEAKDCSQKFTTKLKDSHAELGIQSVQDDNAKKAILKTHGTKDKGAKELEELFKSLESLS 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 KEAKDCSQKFTTKLKDSHAELGIQSVQDDNAKKAILKTHGTKDKGAKELEELFKSLESLS 540

Qy  541 KAAQAALTNSVKELTNPVVA 560
          ||||||||||||||||||||
Db  541 KAAQAALTNSVKELTNPVVA 560
```

```
US-09-596-746-52
; Sequence 52, Application US/0959674A
; GENERAL INFORMATION:
;  APPLICANT: Dattwyler, Raymond J.
;  APPLICANT: Seinost, Gerald
;  APPLICANT: Dykhuizen, Daniel
;  APPLICANT: Luft, Benjamin J.
;  APPLICANT: Maria J.C. Gomes-Solecki
;  TITLE OF INVENTION: Groups of Borrelia burgdorferi and
;  TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
;  FILE REFERENCE: 2631.1002-001
;  CURRENT APPLICATION NUMBER: US/09/596,746A
;  CURRENT FILING DATE: 2000-06-19
;  PRIOR APPLICATION NUMBER: US 60/140,042
;  PRIOR FILING DATE: 1999-06-18
;  NUMBER OF SEQ ID NOS: 84
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 52
;  LENGTH: 560
;  TYPE: PRT
;  ORGANISM: ospC Chimera
US-09-596-746-52

  Query Match             100.0%;  Score 2750;  DB 19;  Length 560;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-180;
  Matches  560;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    1 MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAKA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAKA 60

Qy   61 IGKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSEFTN 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 IGKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSEFTN 120
```

```
  Query Match              63.9%;  Score 1758;  DB 19;  Length 410;
  Best Local Similarity    95.1%;  Pred. No. 1.6e-112;
  Matches  366;  Conservative    0;  Mismatches    7;  Indels   12;  Gaps    1;

Qy  188 AESPAMGSNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETLVLSIDE 247
          ||| |   ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   20 AESIGSCSNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETLVLSIDE 79

Qy  248 LAKKAIGQKIDNNNGLAALNNQNGSLLAGAYAISTLITEKLSKLKNLEELKTEIAKAKC 307
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   80 LAKKAIGQKIDNNNGLAALNNQNGSLLAGAYAISTLITEKLSKLKNLEELKTEIAKAKC 139

Qy  308 SEEFTNKLKSGHADLGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGLLKAAQVA 367
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  139 ...
```

```
Db  140 SEEFTNKLKSGHADLGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGLLKAAQVA 199
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy  368 LTNSVKELGH------------RNNSGGDSASTNPDESAKGPNLTVISKKITDSNAFLLA 415
        ||||||||                |||||||||||||||||||||||||||||||||||
Db  200 LTNSVKELTSPVVAESPKKPHMANNSGGDSASTNPDESAKGPNLTVISKKITDSNAFLLA 259

Qy  416 VKEVEALLSSIDELSKAIGKKIKNDGTLDNEANRNESLIAGAYEISKLITQKLSVLNSEE 475
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  260 VKEVEALLSSIDELSKAIGKKIKNDGTLDNEANRNESLIAGAYEISKLITQKLSVLNSEE 319

Qy  476 LKKKIKEAKDCSQKFTTKLKDSHAELGIQSVQDDNAKKAILKTHGTKDKGAKELEELFKS 535
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  320 LKKKIKEAKDCSQKFTTKLKDSHAELGIQSVQDDNAKKAILKTHGTKDKGAKELEELFKS 379

Qy  536 LESLSKAAQAALTNSVKELTNPVVA 560
        |||||||||||||||||||||||||
Db  380 LESLSKAAQAALTNSVKELTNPVVA 404
```

RESULT 4
US-09-596-746A-84
Sequence 84, Application US/0959746A
GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seinost, Gerald
; APPLICANT: Dykhuizen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746A
; CURRENT FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 84
; LENGTH: 411
; TYPE: PRT
; ORGANISM: ospC Chimera
US-09-596-746A-84

Query Match 63.98%; Score 1758; DB 19; Length 411;
Best Local Similarity 95.1%; Pred. No. 1.6e-112;
Matches 366; Conservative 0; Mismatches 7; Indels 12; Gaps 1;

```
Qy  188 AESPAMGSNSGKGGDSASTNPDESAKGPNLTEISKIKITDSNAFVLAVKEVETLVLSIDE 247
        |||  |||||||||||||||||||||||||||  ||||||||||||||||||||||||||
Db  21  AESIGCSNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETLVLSIDE  80

Qy  248 LAKKAIGQKIDNNGLAALNNQNGSLLAGAYAISTLITEKLSKLKNLEELKTEIAKAKKC 307
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  81  LAKKAIGQKIDNNGLAALNNQNGSLLAGAYAISTLITEKLSKLKNLEELKTEIAKAKKC 140

Qy  308 SEEFTNKLKSGHADLGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGLLKAAQVA 367
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  141 SEEFTNKLKSGHADLGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGLLKAAQVA 200

Qy  368 LTNSVKELGH------------RNNSGGDSASTNPDESAKGPNLTVISKKITDSNAFLLA 415
        ||||||||                |||||||||||||||||||||||||||||||||||
Db  201 LTNSVKELTSPVVAESPKKPHMANNSGGDSASTNPDESAKGPNLTVISKKITDSNAFLLA 260

Qy  416 VKEVEALLSSIDELSKAIGKKIKNDGTLDNEANRNESLIAGAYEISKLITQKLSVLNSEE 475
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  261 VKEVEALLSSIDELSKAIGKKIKNDGTLDNEANRNESLIAGAYEISKLITQKLSVLNSEE 320

Qy  476 LKKKIKEAKDCSQKFTTKLKDSHAELGIQSVQDDNAKKAILKTHGTKDKGAKELEELFKS 535
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  321 LKKKIKEAKDCSQKFTTKLKDSHAELGIQSVQDDNAKKAILKTHGTKDKGAKELEELFKS 380
```

```
Qy  536 LESLSKAAQAALTNSVKELTNPVVA 560
        |||||||||||||||||||||||||
Db  381 LESLSKAAQAALTNSVKELTNPVVA 405
```

RESULT 5
US-09-596-746-58
Sequence 58, Application US/0959746
GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seinost, Gerald
; APPLICANT: Dykhuizen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746
; CURRENT FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 58
; LENGTH: 386
; TYPE: PRT
; ORGANISM: ospC Chimera
US-09-596-746-58

Query Match 63.88%; Score 1754; DB 19; Length 386;
Best Local Similarity 95.8%; Pred. No. 2.8e-112;
Matches 364; Conservative 0; Mismatches 4; Indels 12; Gaps 1;

```
Qy  193 MGSNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETLVLSIDELAKKA 252
        | |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1  MCSNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETLVLSIDELAKKA  60

Qy  253 IGQKIDNNGLAALNNQNGSLLAGAYAISTLITEKLSKLKNLEELKTEIAKAKKCSEEFT 312
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  61  IGQKIDNNGLAALNNQNGSLLAGAYAISTLITEKLSKLKNLEELKTEIAKAKKCSEEFT 120

Qy  313 NKLKSGHADLGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGLLKAAQVALTNSV 372
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 NKLKSGHADLGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGLLKAAQVALTNSV 180

Qy  373 KELGH------------RNNSGGDSASTNPDESAKGPNLTVISKKITDSNAFLLAVKEVE 420
        |||                |||||||||||||||||||||||||||||||||||||||||
Db  181 KELTSPVVAESPKKPHMANNSGGDSASTNPDESAKGPNLTVISKKITDSNAFLLAVKEVE 240

Qy  421 ALLSSIDELSKAIGKKIKNDGTLDNEANRNESLIAGAYEISKLITQKLSVLNSEELKKKI 480
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 ALLSSIDELSKAIGKKIKNDGTLDNEANRNESLIAGAYEISKLITQKLSVLNSEELKKKI 300

Qy  481 KEAKDCSQKFTTKLKDSHAELGIQSVQDDNAKKAILKTHGTKDKGAKELEELFKSLESLS 540
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 KEAKDCSQKFTTKLKDSHAELGIQSVQDDNAKKAILKTHGTKDKGAKELEELFKSLESLS 360

Qy  541 KAAQAALTNSVKELTNPVVA 560
        ||||||||||||||||||||
Db  361 KAAQAALTNSVKELTNPVVA 380
```

RESULT 6
US-09-596-746A-58
Sequence 58, Application US/0959746A
GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seinost, Gerald
; APPLICANT: Dykhuizen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and

RESULT 7

; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746A
; CURRENT FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 58
; LENGTH: 386
; TYPE: PRT
; ORGANISM: ospC Chimera
US-09-596-746a-58

Query Match   63.8%; Score 1754; DB 19; Length 386;
Best Local Similarity   95.8%; Pred. No. 2.8e-112;
Matches 364; Conservative   0; Mismatches   4; Indels 12; Gaps   1;

```
Qy  193 MGSNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVEILVLSIDELAKKA 252
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MCSNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVEILVLSIDELAKKA 60

Qy  253 IGQKIDNNNGLAALNNQNGSLLAGAYAISTLLITEKLSKLKNLEELKTEIAKAKKCSEEFT 312
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 IGQKIDNNNGLAALNNQNGSLLAGAYAISTLLITEKLSKLKNLEELKTEIAKAKKCSEEFT 120

Qy  313 NKLKSGHADLGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGLLKAAQVALTNSV 372
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 NKLKSGHADLGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGLLKAAQVALTNSV 180

Qy  373 KELGH----------------RNNSGGDSASTNPDESAKGPNLTVISKKITDSNAFLLAVKEVE 420
             ||                 ||||||||||||||||||||||||||||||||||
Db  181 KELTSPVVAESPKKPHMANNSGGDSASTNPDESAKGPNLTVISKKITDSNAFLLAVKEVE 240

Qy  421 ALLSSIDELSKAIGKKIKNDGTLDNEANRNESLIAGAYEISKLITQKLSVLNSEELKKKI 480
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 ALLSSIDELSKAIGKKIKNDGTLDNEANRNESLIAGAYEISKLITQKLSVLNSEELKKKI 300

Qy  481 KEAKDCSQKFTTKLKDSHAELGIQSVQDDNAKKAILKTHGTKDKGAKELEELFKSLESLS 540
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 KEAKDCSQKFTTKLKDSHAELGIQSVQDDNAKKAILKTHGTKDKGAKELEELFKSLESLS 360

Qy  541 KAAQAALTNSVKELTNPVVA 560
        ||||||||||||||||||||
Db  361 KAAQAALTNSVKELTNPVVA 380
```

RESULT 7
US-09-596-746A-56
Sequence 56, Application US/0959674GA
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seinost, Gerald
; APPLICANT: Dykhuizen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia afzelii That Cause Lyme Disease in Humans
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746A
; CURRENT FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 56
; LENGTH: 385
; TYPE: PRT
; ORGANISM: ospC Chimera
US-09-596-746A-56

RESULT 8
US-09-596-746-56
Sequence 56, Application US/09596746
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seinost, Gerald
; APPLICANT: Dykhuizen, Daniel
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
;   Borrelia afzelii That Cause Lyme Disease in Humans
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746
; CURRENT FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 56
; LENGTH: 384
; TYPE: PRT
; ORGANISM: ospC Chimera
US-09-596-746-56

Query Match   63.1%; Score 1734.5; DB 19; Length 385;
Best Local Similarity   67.0%; Pred. No. 6.1e-111;
Matches 375; Conservative   1; Mismatches   3; Indels 181; Gaps   1;

```
Qy    1 MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA 60

Qy   61 IGKKIHQNNGLDTEYNHNGSLLAGAYAISTLLIKQKLDGLKNEGLKEKIDAAKKCSEFTN 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 IGKKIHQNNGLDTEYNHNGSLLAGAYAISTLLIKQKLDGLKNEGLKEKIDAAKKCSEFTN 120

Qy  121 KLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 KLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 180

Qy  181 ELITSPVVAESPAMGSNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVET 240
        ||||||||||||
Db  181 ELITSPVVAESP---------------------------------------------- 192

Qy  241 LVLSIDELAKKAIGQKIDNNNGLAALNNQNGSLLAGAYAISTLLITEKLSKLKNLEELKTE 300

Db  192 ---------------------------------------------------------- 191

Qy  301 IAKAKKCSEEFTNKLKSGHADLGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGL 360

Db  192 ---------------------------------------------------------- 191

Qy  361 LKAAQVALTNSVKELGHRNNSGGDSASTNPDESAKGPNLTVISKKITDSNAFLLAVKEVE 420
            |:  ||||||||||||||||||||||||||||||||||||||||
Db  192 -----------KKPHGNNSGGDSASTNPDESAKGPNLTVISKKITDSNAFLLAVKEVE 239

Qy  421 ALLSSIDELSKAIGKKIKNDGTLDNEANRNESLIAGAYEISKLITQKLSVLNSEELKKKI 480
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  240 ALLSSIDELSKAIGKKIKNDGTLDNEANRNESLIAGAYEISKLITQKLSVLNSEELKKKI 299

Qy  481 KEAKDCSQKFTTKLKDSHAELGIQSVQDDNAKKAILKTHGTKDKGAKELEELFKSLESLS 540
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  300 KEAKDCSQKFTTKLKDSHAELGIQSVQDDNAKKAILKTHGTKDKGAKELEELFKSLESLS 359

Qy  541 KAAQAALTNSVKELTNPVVA 560
        ||||||||||||||||||||
Db  360 KAAQAALTNSVKELTNPVVA 379
```

```
RESULT 9
US-09-596-746-78
Sequence 78, Application US/0959746
GENERAL INFORMATION:
    APPLICANT:  Dattwyler, Raymond J.
    APPLICANT:  Seinost, Gerald
    APPLICANT:  Dykhuizen, Daniel
    APPLICANT:  Luft, Benjamin J.
    APPLICANT:  Maria J.C. Gomes-Solecki
    TITLE OF INVENTION:  Groups of Borrelia burgdorferi and
    TITLE OF INVENTION:  Borrelia afzelii That Cause Lyme Disease in Humans
    FILE REFERENCE:  2631.1002-001
    CURRENT APPLICATION NUMBER:  US/09/596,746
    CURRENT FILING DATE:  2000-06-16
    PRIOR APPLICATION NUMBER:  US 60/140,042
    PRIOR FILING DATE:  1999-06-18
    NUMBER OF SEQ ID NOS:  84
    SOFTWARE:  FastSEQ for Windows Version 4.0
    SEQ ID NO 78
    LENGTH: 408
    TYPE: PRT
    ORGANISM: ospC Chimera
US-09-596-746-78

Query Match       62.8%;  Score 1726.5;  DB 19;  Length 408;
Best Local Similarity    66.7%;  Pred. No. 2.4e-110;
Matches  373;  Conservative  2;  Mismatches  3;  Indels 181;  Gaps  1;

Qy    2 ACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAI 61
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 ACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAI 60

Qy   62 GKKIHQNNGLDTEYNHNGSLLAGAYAISTLIQKLDGLKNEGLKEKIDAAKKCSETFTNK 121
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 GKKIHQNNGLDTEYNHNGSLLAGAYAISTLIQKLDGLKNEGLKEKIDAAKKCSETFTNK 120

Qy  122 LKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 181
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 LKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 180

Qy  182 LTSPVVAESPAMGSNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETL 241
          ||||||||||
Db  181 LTSPVVAESP------------------------------------------------- 190

Qy  242 VLSIDELAKKAIGQKIDNNNGLLAALNNQNGSLLAGAYAISTLITEKLSKLKNLEELKTEI 301

Db  191 ---------------------------------------------------------- 190

Qy  302 AKAKKCSEEFTNKLKSGHADLGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGLL 361

Db  191 ---------------------------------------------------------- 190

Qy  362 KAAQVALTNSVKELGHRNNSGGDSASTNPDESARGPNLTVISKKITDSNAFLLAVKEVEA 421
              |.| ||||||||||||||||||||||||||||||||||||||||
Db  191 -----KKPFHGHNSGGGDSASTNPDESAKGPNLTVISKKITDSNAFLLAVKEVEA 239

Qy  422 LLSSIDELSKAIGKKIKNDGTLDNEANRNESLIAGAYEISKLITQKLSVLNSEELKKKIK 481
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  240 LLSSIDELSKAIGKKIKNDGTLDNEANRNESLIAGAYEISKLITQKLSVLNSEELKKKIK 299

Qy  482 EAKDCSQKFTTKLKDSHAELGIQSVQODNAKKAILKTHGTKDKGAKELEELFKSLESLSK 541
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  300 EAKDCSQKFTTKLKDSHAELGIQSVQODNAKKAILKTHGTKDKGAKELEELFKSLESLSK 359

Qy  542 AAQAALTNSVKELTNPVVA 560
          |||||||||||||||||||
Db  360 AAQAALTNSVKELTNPVVA 378
```

```
RESULT 10
US-09-596-746A-78
Sequence 78, Application US/0959746A
GENERAL INFORMATION:
    APPLICANT:  Dattwyler, Raymond J.
    APPLICANT:  Seinost, Gerald
    APPLICANT:  Dykhuizen, Daniel
    APPLICANT:  Luft, Benjamin J.
    APPLICANT:  Maria J.C. Gomes-Solecki
    TITLE OF INVENTION:  Groups of Borrelia burgdorferi and
    TITLE OF INVENTION:  Borrelia afzelii That Cause Lyme Disease in Humans
    FILE REFERENCE:  2631.1002-001
    CURRENT APPLICATION NUMBER:  US/09/596,746A
    CURRENT FILING DATE:  2000-06-19
    PRIOR APPLICATION NUMBER:  US 60/140,042
    PRIOR FILING DATE:  1999-06-18
    NUMBER OF SEQ ID NOS:  84
    SOFTWARE:  FastSEQ for Windows Version 4.0
    SEQ ID NO 78
    LENGTH: 409
    TYPE: PRT
    ORGANISM: ospC Chimera
US-09-596-746A-78

Query Match       62.8%;  Score 1726.5;  DB 19;  Length 409;
Best Local Similarity    66.7%;  Pred. No. 2.4e-110;
Matches  373;  Conservative  2;  Mismatches  3;  Indels 181;  Gaps  1;

Qy    2 ACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAI 61
          :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   25 SCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAI 84

Qy   62 GKKIHQNNGLDTEYNHNGSLLAGAYAISTLIQKLDGLKNEGLKEKIDAAKKCSETFTNK 121
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   85 GKKIHQNNGLDTEYNHNGSLLAGAYAISTLIQKLDGLKNEGLKEKIDAAKKCSETFTNK 144

Qy  122 LKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 181
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  145 LKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 204

Qy  182 LTSPVVAESPAMGSNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETL 241
          ||||||||||
Db  205 LTSPVVAESP------------------------------------------------- 214

Qy  242 VLSIDELAKKAIGQKIDNNNGLLAALNNQNGSLLAGAYAISTLITEKLSKLKNLEELKTEI 301

Db  215 ---------------------------------------------------------- 214

Qy  302 AKAKKCSEEFTNKLKSGHADLGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGLL 361

Db  215 ---------------------------------------------------------- 214

Qy  362 KAAQVALTNSVKELGHRNNSGGDSASTNPDESARGPNLTVISKKITDSNAFLLAVKEVEA 421
              |.| ||||||||||||||||||||||||||||||||||||||||
Db  215 -----KKPFHGHNSGGGDSASTNPDESAKGPNLTVISKKITDSNAFLLAVKEVEA 263

Qy  422 LLSSIDELSKAIGKKIKNDGTLDNEANRNESLIAGAYEISKLITQKLSVLNSEELKKKIK 481
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  264 LLSSIDELSKAIGKKIKNDGTLDNEANRNESLIAGAYEISKLITQKLSVLNSEELKKKIK 323

Qy  482 EAKDCSQKFTTKLKDSHAELGIQSVQODNAKKAILKTHGTKDKGAKELEELFKSLESLSK 541
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  324 EAKDCSQKFTTKLKDSHAELGIQSVQODNAKKAILKTHGTKDKGAKELEELFKSLESLSK 383

Qy  542 AAQAALTNSVKELTNPVVA 560
          |||||||||||||||||||
Db  384 AAQAALTNSVKELTNPVVA 402
```

```
Qy    2 ACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAI 61
          :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2 ACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAI 61
```

```
Qy    2 ACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAI 61
          :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   26 SCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAI 85

Qy   62 GKKIHQNNGLDTEYNHNGSLLAGAYAISTLIQKLDGLKNEGLKEKIDAAKKCSETFTNK 121
```

```
Db     86  GKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSEFTNK  145
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    122  LKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE  181
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    146  LKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE  205
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    182  LTSPVVAESPAMGSNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETL  241
           ||||||||
Db    206  LTSPVVAESP----------------------------------------------     215
Qy    242  VLSIDELAKKAIGQKIDNNNGLAALNNQNGSLLAGAYAYAISTLITEKLSKLNLEELKTEI  301
Db    216  ----------------------------------------------------------  215
Qy    302  AKAKKCSEEFTNKLKSGHADLGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGLL  361
Db    216  ----------------------------------------------------------  215
Qy    362  KAAQVALTNSVKELGHRNNSGGDSASTNPDESAKGPNLTVISKKITDSNAFLLAVKEVEA  421
           |:  |||||||||||||||
Db    385  ------------KKPFHGNNSGGDSASTNPDESAKGPNLTVISKKITDSNAFLLAVKEVEA  264
Qy    542  AAQAALTNSVKELTNPVVA  560
           ||||||||||||:||||||
Db    385  AAQAALTNSVKELTNPVVA  403
```

```
RESULT 11
US-09-596-746A-26
Sequence 26, Application US/0959674A6A-26
GENERAL INFORMATION:
;    APPLICANT:  Dattwyler, Raymond J.
;    APPLICANT:  Seinost, Gerald
;    APPLICANT:  Dykhuizen, Danial
;    APPLICANT:  Luft, Benjamin J.
;    APPLICANT:  Maria J.C. Gomes-Solecki
;    TITLE OF INVENTION:  Groups of Borrelia burgdorferi and
;    TITLE OF INVENTION:  Borrelia afzelii That Cause Lyme Disease in Humans
;    FILE REFERENCE:  2631.1002-001
;    CURRENT APPLICATION NUMBER:  2000-06-19
;    CURRENT FILING DATE:  2000-06-19
;    PRIOR APPLICATION NUMBER:  US 60/140,042
;    PRIOR FILING DATE:  1999-06-18
;    NUMBER OF SEQ ID NOS:  84
;    SOFTWARE:  FastSEQ for Windows Version 4.0
;    SEQ ID NO 26
;    LENGTH:  374
;    TYPE:  PRT
;    ORGANISM:  Artificial Sequence
;    FEATURE:
;    OTHER INFORMATION:  OspC Chimera
US-09-596-746A-26

Query Match             54.4%;
Best Local Similarity   82.8%;     Pred. No. 1.5e-94;
Matches  310;  Conservative  21;  Mismatches  43;  Indels   1;  Gaps    1;
```

```
Qy     1  MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA  60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1  MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA  60
Qy    61  IGKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN  120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    61  IGKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN  120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   121  KLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK  180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  KLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK  180
Qy   181  ELTSPVVAESPAMGSNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVET  240
          |||||||      :||:||    ||| | |||||||||||||||||||||||||||||||
Db   181  ELTSPVVAESPAMVNNSGKDGNT-SANSADESVKGPNLTEISKKITESNAVLLAVKEIEI  239
Qy   241  LVLSIDELAKKAIGQKIDNNNGLAALNNQNGSLLAGAYAYAISTLITEKLSKLNLEELKTE  300
          ||:  ||||       |||    |||||||||||||||||||:|||||::|||:|::  :
Db   240  LLASIDELATKAIGKKIQQNGGLAVEAGHNGTLLAGAYTISKLLTQKLDGLKNSEKLKEK  299
Qy   301  IAKAKKCSEEFTNKLKSGHADLGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGL  360
          ||    | |  | |  ||  :|| ||    | | |:|  || |: ||   ||||||||  |
Db   300  IENAKKCSEDFTKKLEGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENL  359
Qy   361  LKAAQVALTNSVKEL  375
          |||: | |||||||
Db   360  AKAAKEMLANSVKEL  374
```

```
RESULT 12
US-09-596-746A-24
Sequence 24, Application US/0959674A6A
GENERAL INFORMATION:
;    APPLICANT:  Dattwyler, Raymond J.
;    APPLICANT:  Seinost, Gerald
;    APPLICANT:  Dykhuizen, Danial
;    APPLICANT:  Luft, Benjamin J.
;    APPLICANT:  Maria J.C. Gomes-Solecki
;    TITLE OF INVENTION:  Groups of Borrelia burgdorferi and
;    TITLE OF INVENTION:  Borrelia afzelii That Cause Lyme Disease in Humans
;    FILE REFERENCE:  2631.1002-001
;    CURRENT APPLICATION NUMBER:  US/09/596,746A
;    CURRENT FILING DATE:  2000-06-19
;    PRIOR APPLICATION NUMBER:  US 60/140,042
;    PRIOR FILING DATE:  1999-06-18
;    NUMBER OF SEQ ID NOS:  84
;    SOFTWARE:  FastSEQ for Windows Version 4.0
;    SEQ ID NO 24
;    LENGTH:  375
;    TYPE:  PRT
;    ORGANISM:  Artificial Sequence
;    FEATURE:
;    OTHER INFORMATION:  OspC Chimera
US-09-596-746A-24

Query Match             54.3%;
Best Local Similarity   83.5%;     Pred. No. 3e-94;
Matches  313;  Conservative  20;  Mismatches  40;  Indels   2;  Gaps    2;
```

```
Qy     1  MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA  60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1  MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA  60
Qy    61  IGKKIHQNNGLDTEYNINGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSEFTTN  120
          ||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db    61  IGKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN  120
Qy   121  KLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK  180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  KLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK  180
Qy   181  ELTSPVVAESPAMGSNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVET  240
          |||||||      |:: |||   ||| | |||||||||||||||||||||||||||||||
Db   181  ELTSPVVAESPAMVNNSGKDGNT-SANSADESVKGPNLTEISKKITESNAVLLAVKEVET  239
Qy   241  LVLSIDELAKKAIGQKIDNNNGLAALNNQNGSLLAGAYAISTLITEKLSKLNLEELKTE  300
          |:: ||||  :||::|| | |||||||:::::|||||||:|||:|||:|:|:|:|||
Db   240  LLTSIDELA-KAIGKKINDVSLDNEADHNGSLISGAYLISNLITKKISAIKDSGELKAE  298
```

```
Qy   301 IAKAKKCSEEFTNKLKSGHADLGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGL 360
         |||:|  |||||  :| ||||| |  ||||  |:|||||: |
Db   299 IEKAKKCSEEFTAKLKGEHTDLGKEGVTDDNAKKAILKTNNDKTKGADELEKLFESVKNL 358

Qy   361 LKAAQVALTNSVKEL 375
         |||:  ||:||||||
Db   359 SKAAKEMLTNSVKEL 373
```

RESULT 13
US-09-596-746-26
; Sequence 26, Application US/09596746
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seinost, Gerald
; APPLICANT: Dykhuizen, Danial
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746
; CURRENT FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 26
; LENGTH: 373
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: OspC Chimera
US-09-596-746-26

  Query Match   54.2%; Score 1491.5; DB 19; Length 373;
  Best Local Similarity   82.6%; Pred. No. 3.2e-94;
  Matches   309; Conservative   21; Mismatches   43; Indels   1; Gaps   1;

```
Qy     2 ACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAI 61
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 ACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAI 60

Qy    62 GKKIHQNNGLDTEYNHNGSLLAGAYAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNK 121
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 GKKIHQNNGLDTEYNHNGSLLAGAYAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNK 120

Qy   122 LKEKHTDLGKEGVTDADADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 181
         |||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db   121 LKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 180

Qy   182 LTSPVVAESPAMGSNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETL 241
         |||||||||||: .|  |  :|  | ||.||.||||||||||||:.|| ||||||:.||
Db   181 LTSPVVAESPAMVVNNSGKDGNT-SANSADESVKGPNLTEISKKITESNAVVLAVKEIETL 239

Qy   242 VLSIDELAKKAIGQKIDNNNGLAALNNQNGSLLAGAYAISTLITEKLSKLKNLEELKTEI 301
         : ||||||:|||| ||  ||||.:  ||||.||||||:||||||:|||.|||:||||:|
Db   240 LASIDELATKAIGKKIQQNGGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKEKI 299

Qy   302 AKAKKCSEEFTNKLKSGHADLGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGLL 361
         .|||||||:||.||.:||::||.:.|.|::|||||| |: |.|||| ||:|||.:||.||
Db   300 ENAKKCSEDFTKKLEGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLA 359

Qy   362 KAAQVALTNSVKEL 375
         |||:  ||.||||||
Db   360 KAAKEMLANSVKEL 373
```

RESULT 14
US-09-596-746-62
; Sequence 62, Application US/0959674

RESULT 15
US-09-596-746A-62
; Sequence 62, Application US/09-596-746A-62
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Seinost, Gerald
; APPLICANT: Dykhuizen, Danial
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Maria J.C. Gomes-Solecki
; TITLE OF INVENTION: Groups of Borrelia burgdorferi and
; TITLE OF INVENTION: Borrelia afzelii That Cause Lyme Disease in Humans
; FILE REFERENCE: 2631.1002-001
; CURRENT APPLICATION NUMBER: US/09/596,746A
; CURRENT FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: US 60/140,042
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 62
; LENGTH: 397
; TYPE: PRT
; ORGANISM: ospc Chimera
US-09-596-746-62

  Query Match   54.1%; Score 1488.5; DB 19; Length 397;
  Best Local Similarity   82.4%; Pred. No. 5.7e-94;
  Matches   308; Conservative   22; Mismatches   43; Indels   1; Gaps   1;

```
Qy     2 ACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAI 61
         :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    25 SCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAI 84

Qy    62 GKKIHQNNGLDTEYNHNGSLLAGAYAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNK 121
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    85 GKKIHQNNGLDTEYNHNGSLLAGAYAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNK 144

Qy   122 LKEKHTDLGKEGVTDADADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 181
         |||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db   145 LKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 204

Qy   182 LTSPVVAESPAMGSNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETL 241
         |||||||||||: .|  |  :|  | ||.||.||||||||||||:.|| ||||||:.||
Db   205 LTSPVVAESPAMVVNNSGKDGNT-SANSADESVKGPNLTEISKKITESNAVVLAVKEIETL 263

Qy   242 VLSIDELAKKAIGQKIDNNNGLAALNNQNGSLLAGAYAISTLITEKLSKLKNLEELKTEI 301
         : ||||||:|||| ||  ||||.:  ||||.||||||:||||||:|||.|||:||||:|
Db   264 LASIDELATKAIGKKIQQNGGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKEKI 323

Qy   302 AKAKKCSEEFTNKLKSGHADLGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGLL 361
         .|||||||:||.||.:||::||.:.|.|::|||||| |: |.|||| ||:|||.:||.||
Db   324 ENAKKCSEDFTKKLEGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLA 383

Qy   362 KAAQVALTNSVKEL 375
         |||:  ||.||||||
Db   384 KAAKEMLANSVKEL 397
```

```
;   SEQ ID NO 62
;   LENGTH: 398
;   TYPE: PRT
;   ORGANISM: ospC Chimera
US-09-596-746a-62

Query Match         54.1%;  Score 1488.5;  DB 19;  Length 398;
Best Local Similarity   82.4%;  Pred. No. 5.8e-94;
Matches  308;  Conservative   22;  Mismatches   43;  Indels    1;  Gaps    1;

Qy        2 ACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAI 61
            :||||:|||||||||||||||||||||||||||||||||||||||||||:|||||||||
Db       26 SCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAI 85

Qy       62 GKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNK 121
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       86 GKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNK 145

Qy      122 LKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 181
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      146 LKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKE 205

Qy      182 LTSPVVAESPAMGSNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETL 241
            ||||||||||||:||::|:|:|::||::|||||||||||||||||||:|:||:|||||
Db      206 LTSPVVAESPAMVNNSGKDGNT-SANSADESVKGPNLTEISKKITESNAVVLAVKEIETL 264

Qy      242 VLSIDELAKKAIGQKIDNNNGLAALNNQNGSLLAGAYAISTLITEKLSKLKNLEELKTEI 301
            :||||||||||||::|:||  ||:||:||||||||||||||||::|||||:|:|||::
Db      265 LASIDELATKAIGKKIQQNGGLAVEAGHNGTLLAGAYTISKLITQKLDGLKNSEKLKEKI 324

Qy      302 AKAKKCSEEFTNKLKSGHADLGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGLL 361
            |||||||||||||:||  ||::|   |::||:||| |  |||| | |:|  ||:::|||
Db      325 ENAKKCSEDFTKKLEGEHAQLGIENVTDENAKKAILITDAAKDKGAAELEKLFKAVENLA 384

Qy      362 KAAQVALTNSVKEL 375
            |||: | ||||||
Db      385 KAAKEMLANSVKEL 398
```


Search completed: March 18, 2002, 10:08:47
Job time: 975 sec

OM protein - protein search, using sw model

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

Run on: March 18, 2002, 09:55:35 ; Search time 55.5 Seconds
                                              (without alignments)
                                              227.060 Million cell updates/sec

Title: OM protein
Perfect score: 2750
Sequence: 1 MACNNSGKDGNTSANSADES..........KAAQAALTNSVKELTNPVVA 560
Scoring table: BLOSUM62
                  Gapop 10.0 , Gapext 0.5
Searched: 212252 seqs, 22503292 residues

Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries

Database :   Issued_Patents_AA.*
                  1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
                  2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
                  3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
                  4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
                  5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
                  6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Maximum DB seq length: 200000000
Minimum DB seq length: 0
1 number of hits satisfying chosen parameters: 212252

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	986.5	35.9	466	4	US-08-235-836C-110	Sequence 110, App
2	932	33.9	210	4	US-08-235-836C-30	Sequence 30, Appl
3	928	33.7	209	4	US-09-196-293-15	Sequence 15, Appl
4	928	33.7	210	1	US-08-158-353-3	Sequence 3, Appl
5	922	33.5	588	5	US-08-209-603E-15	Sequence 122, App
6	921	33.5	210	4	US-08-235-836C-122	Sequence 122, Appl
7	915	33.3	466	4	US-08-209-603E-15	Sequence 15, Appl
8	889	32.3	210	4	US-08-235-836C-107	Sequence 107, App
9	886	32.2	212	1	US-08-235-836C-34	Sequence 34, Appl
10	882	32.1	207	4	US-08-158-353-4	Sequence 4, Appl
11	858	31.2	212	4	US-08-235-836C-36	Sequence 36, Appl
12	858	31.2	212	4	US-09-196-293-11	Sequence 11, Appl
13	676	24.6	212	1	US-08-209-603E-11	Sequence 11, Appl
14	661	24.0	212	4	US-08-158-353-2	Sequence 2, Appl
15	661	24.0	212	4	US-08-031-295-2	Sequence 2, Appl
16	616.5	22.4	209	4	US-07-903-580-2	Sequence 2, Appl
17	188	6.8	2285	5	US-08-235-836C-32	Sequence 32, Appl
18	187	6.8	1164	4	US-09-308-375-2	Sequence 2, Appl
19	186	6.8	1786	4	US-08-923-992A-2	Sequence 2, Appl
20	184	6.7	1164	4	US-08-973-462-8	Sequence 8, Appl
21	1588	6.7	1164	5	PCT-US93-07261-11	Sequence 11, Appl
22	182.5	6.6	1588	5	PCT-US93-07261-11	Sequence 11, App
23	182.5	6.6	1663	5	PCT-US93-07261-16	Sequence 16, Appl
24	181	6.6	630	4	US-08-973-462-9	Sequence 9, Appl
25	181	6.6	688	3	US-09-141-047-8	Sequence 8, Appl
26	179	6.5	1128	4	US-08-923-992A-6	Sequence 6, Appl
27	177.5	6.5	1098	4	US-08-923-992A-8	Sequence 8, Appl
28	177.5	6.5	1196	1	US-08-144-121-4	Sequence 4, Appl
29	177.5	6.5	1196	2	US-08-735-893-4	Sequence 4, Appl
30	177	6.4	1388	1	US-08-685-576-1	Sequence 1, Appl
31	176.5	6.4	3248	1	US-08-353-700-1	Sequence 1, Appl
32	176.5	6.4	3248	5	PCT-US95-16216-1	Sequence 1, Appl
33	175.5	6.4	2482	5	US-08-328-254-6	Sequence 6, Appl
34	174	6.3	740	1	US-08-257-073-5	Sequence 5, Appl
35	173	6.3	1038	4	US-09-541-782-4	Sequence 4, Appl
36	171	6.2	1104	4	US-08-923-992A-4	Sequence 4, Appl
37	171	6.2	1388	2	US-08-685-576-4	Sequence 4, Appl
38	170.5	6.2	1151	3	US-08-840-006-5	Sequence 5, Appl
39	170.5	6.2	1200	2	US-08-840-006-6	Sequence 6, Appl
40	170.5	6.2	3111	2	US-08-460-309-4	Sequence 4, Appl
41	167	6.1	3111	2	US-08-125-077-4	Sequence 4, Appl
42	167	6.1	1147	3	US-08-470-260-5	Sequence 5, Appl
43	167	6.1	1147	4	US-08-471-491-5	Sequence 5, Appl
44	167	6.1	2101	5	US-08-195-487-4	Sequence 4, Appl
45	167	6.1	2101	5	PCT-US93-06160-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1
US-08-235-836C-110
; Sequence 110, Application US/0823 5836C
; Patent No. 6248562
; GENERAL INFORMATION:
;   APPLICANT: Dunn, John J.
;   APPLICANT: Luft, Benjamin J.
;   TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising
;   TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor
;   NUMBER OF SEQUENCES: 144
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Brookhaven National Laboratory
;   STREET:
;   CITY: Upton
;   STATE: NY
;   COUNTRY: USA
;   ZIP: 11973
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: Patentin Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/235,836C
;   FILING DATE: 29-APR-1994
;   CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 08/148,191
;   FILING DATE: 01-11-93
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Bogosian, Margaret C.
;   REGISTRATION NUMBER: 25,324
;   REFERENCE/DOCKET NUMBER: BNL93-28A
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (516) 282-7338
;   TELEFAX: (516) 282-3729
;   INFORMATION FOR SEQ ID NO: 110:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 466 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-08-235-836C-110

Query Match   35.9%; Score 986.5; DB 4; Length 466;
Best Local Similarity   51.1%; Pred. No. 8e-63;
Matches 256; Conservative 42; Mismatches 132; Indels 71; Gaps 17;

Qy   1 MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA 60

RESULT 2
US-08-235-836C-30
Sequence 30, Application US/08235836C
Patent No. 6248562
GENERAL INFORMATION:
  APPLICANT: Dunn, John J.
  APPLICANT: Luft, Benjamin J.
  TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising
  TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor
  NUMBER OF SEQUENCES: 144
  SOFTWARE: Patentin Release #1.0, Version #1.25
  CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/235,836C
  FILING DATE: 29-APR-1994
  CLASSIFICATION: 435
  PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 08/148,191
  FILING DATE: 01-11-93
  ATTORNEY/AGENT INFORMATION:
  NAME: Bogosian, Margaret C.
  REGISTRATION NUMBER: 25,324
  REFERENCE/DOCKET NUMBER: BNL93-28A
  TELECOMMUNICATION INFORMATION:
  TELEPHONE: (516) 282-7338
  TELEFAX: (516) 282-3729
  CORRESPONDENCE ADDRESS:
  ADDRESSEE: Brookhaven National Laboratory
  STREET:
  CITY: Upton
  STATE: NY
  COUNTRY: USA
  ZIP: 11973
  COMPUTER READABLE FORM:
  MEDIUM TYPE: Floppy disk
  COMPUTER: IBM PC compatible
  OPERATING SYSTEM: PC-DOS/MS-DOS
  SOFTWARE: Patentin Release #1.0, Version #1.25
  CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/235,836C
  FILING DATE: 29-APR-1994
  CLASSIFICATION: 435
  PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 08/148,191
  FILING DATE: 01-11-93
  INFORMATION FOR SEQ ID NO: 30:

```
Qy   61 IGKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   77 IGKKIHQNNGLDTEYNHNGSLLAGRYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN 136

Qy  121 KLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  137 KLKAKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 196

Qy  181 ELTSPVVAESPKPKQNVSSLDEKNSVSVD--LPGEMKVLVSKEKNKDGKY------D 246
         ||||||||
Db  197 ELTSPVVAESPAMGSNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVET 240

Qy  247 LIATVDKLELKGTS---DKNNGSGVLE----GVKADKSKVKLTISDDLGQ-TTLEVFKED 298
Db  241 LVLSIDELAKKAIGQKIDNNNGLAALNNQNGSLLAGAYAISTLITEKLSKLKNLEELKTE 300

Qy  299 GKTLVSKKVTSKDKSSTEEKFNEKGEVSEKITTRADGTR------LEYTGIKSDGSGKA 351
Db  301 ------------IAKAKKCSEE-FTNKLKSGHADLGKQDATDDHAKKAAILKTHATTDKGAKEF 350

Qy  352 KEVLKGYVLEGLITAEKTTLV--VKE-------GTVTLSKNTSKSGE----VSVELNDTD 398
Db  351 KDLFES--VEGLLKAAQVALTNSVKELGHRNNSGGDSASTNPDESAKGPNLTVISKKITD 408

Qy  409 SNAFLLAVKEVEALLSSIDELS-KAIGKKIKN-----DGTLDNEA--NRNESLIAGAYEI 460
Db  409 ------------------------------------
Db  399 SSA---ATKKTAAWNDSTSTLTISADSKKTKDLVFLTDGTITVVQQYNTAGTISLEGSASEI 455

Qy  461 SKLITQKLSVLNSEELKKKIK 481
         | | |   | :|
Db  456 K----------NLSELKNALK 466
```

```
SEQUENCE CHARACTERISTICS:
  LENGTH: 210 amino acids
  TYPE: amino acid
  TOPOLOGY: linear
  MOLECULE TYPE: protein
US-08-235-836C-30

Query Match             33.9%;   Score 932;  DB 4;  Length 210;
Best Local Similarity   98.4%;   Pred. No. 2.1e-59;
Matches  188;  Conservative  2;  Mismatches  1;  Indels  0;  Gaps  0;

Qy    1 MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAARA 60
          ::||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   17 ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAARA 76

Qy   61 IGKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   77 IGKKIHQNNGLDTEYNHNGSLLAGRYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN 136

Qy  121 KLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  137 KLKAKHTDLGKEGVTDADAKEAILKINGTKIKGAEELGKLFESVEVLSKAAKEMLANSVK 196

Qy  181 ELTSPVVAESP 191
         |||||||||||
Db  197 ELTSPVVAESP 207
```

RESULT 3
US-09-196-293-15
Sequence 15, Application US/09196293
Patent No. 6183755
GENERAL INFORMATION:
  APPLICANT: Fuchs, Renate
  APPLICANT: Motz, Manfred
  APPLICANT: Soutscheck, Erwin
  APPLICANT: Wilske, Bettina
  APPLICANT: Preac-Mursic, Vera
  TITLE OF INVENTION: Active proteins from Borrelia
  FILE REFERENCE: 738.001US2
  TITLE OF INVENTION: burgdorferi
  CURRENT APPLICATION NUMBER: US/09/196,293
  CURRENT FILING DATE: 1998-11-19
  EARLIER APPLICATION NUMBER: US 08/209,603
  EARLIER FILING DATE: 1994-03-10
  EARLIER APPLICATION NUMBER: US 07/862,535
  EARLIER FILING DATE: 1992-06-19
  EARLIER APPLICATION NUMBER: WO PCT/EP90/02282
  EARLIER FILING DATE: 1990-12-21
  EARLIER APPLICATION NUMBER: DE P39 42 728.5
  EARLIER FILING DATE: 1989-12-22
  EARLIER APPLICATION NUMBER: DE P40 18 988.0
  EARLIER FILING DATE: 1990-06-13
  SOFTWARE: FastSEQ for Windows Version 4.0
  NUMBER OF SEQ ID NOS: 16
  SEQ ID NO 15
  LENGTH: 209
  TYPE: PRT
  ORGANISM: Borrelia burgdorferi
US-09-196-293-15

```
Query Match             33.7%;   Score 928;  DB 4;  Length 209;
Best Local Similarity   98.4%;   Pred. No. 4e-59;
Matches  188;  Conservative  2;  Mismatches  1;  Indels  0;  Gaps  0;

Qy    1 MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAARA 60
          ::|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   17 ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAARA 76

Qy   61 IGKKIHQNNGLDTEYNINGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db   77 IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAARKCSETFTN 136
Qy  121 KLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 180
Db  137 KLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 196
Qy  181 ELTSPVVAESP 191
         |||||||||||
Db  197 ELTSPVVAESP 207
```

RESULT 4
US-08-158-353-3
Sequence 3, Application US/08158353
Patent No. 5520862
GENERAL INFORMATION:
APPLICANT: Padula, Steven J.
TITLE OF INVENTION: Methods for Diagnosing Early Lyme
TITLE OF INVENTION: Disease
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Militia Drive
CITY: Lexington
STATE: MA
COUNTRY: USA
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/158,353
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Carroll, Alice O.
REGISTRATION NUMBER: 33,542
REFERENCE/DOCKET NUMBER: UCT93-05
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-861-6240
TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 210 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-158-353-3

Query Match   33.7%;   Score 928; DB 1; Length 210;
Best Local Similarity 98.4%; Pred. No. 4e-59;
Matches 188; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```
Qy    1 MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA 60
         ::||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   17 ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA 76

Qy   61 IGKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAARKCSETFTN 120
         |||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db   77 IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAARKCSETFTN 136

Qy  121 KLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  137 KLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 196

Qy  181 ELTSPVVAESP 191
         |||||||||||
Db  197 ELTSPVVAESP 207
```

RESULT 5
US-08-235-836C-122
Sequence 122, Application US/08235836C
Patent No. 6248562
GENERAL INFORMATION:
APPLICANT: Dunn, John J.
APPLICANT: Luft, Benjamin J.
TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising
TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor
NUMBER OF SEQUENCES: 144
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brookhaven National Laboratory
STREET:
CITY: Upton
STATE: NY
COUNTRY: USA
ZIP: 11973
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/235,836C
FILING DATE: 29-APR-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/148,191
FILING DATE: 01-11-93
ATTORNEY/AGENT INFORMATION:
NAME: Bogosian, Margaret C.
REGISTRATION NUMBER: 25,324
REFERENCE/DOCKET NUMBER: BNL93-28A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 282-7338
TELEFAX: (516) 282-3729
INFORMATION FOR SEQ ID NO: 122:
SEQUENCE CHARACTERISTICS:
LENGTH: 588 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-235-836C-122

Query Match   33.5%;   Score 922; DB 4; Length 588;
Best Local Similarity 93.6%; Pred. No. 4.3e-58;
Matches 191; Conservative 0; Mismatches 7; Indels 6; Gaps 1;

```
Qy    4 NNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGK 63
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  286 NNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGK 345

Qy   64 KIHQNNGLDTEYNHNGSLLAGRYAISTLIKQKLDGLKNEGLKEKIDAARKCSETFTNKLK 123
         |||||||||||||||||||||  |||||||||||||||||||||||||||||||||||||
Db  346 KIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAARKCSETFTNKLK 405

Qy  124 EKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKELT 183
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  406 EKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKELT 465

Qy  184 SPVVAESPAMGSNSGKGGDSASTN 207
         |||||||    |     |
Db  466 SPVVAESPK-----KPGTMAQYN 483
```

RESULT 6
US-08-209-603E-15
Sequence 15, Application US/08209603E
Patent No. 6248538
GENERAL INFORMATION:
APPLICANT: FUCHS, RENATE

APPLICANT: WILSKE, BETTINA
APPLICANT: PREAC-MURSIC, VERA
APPLICANT: MOTZ, MANFRED
APPLICANT: SOUTSCHECK, ERWIN
TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PROTEINS
TITLE OF INVENTION: FROM BORRELIA BURGDORFERI
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROOKS HAIDT HAFFNER & DELAHUNTY
STREET: 99 PARK AVENUE
CITY: NEW YORK
STATE: NY
COUNTRY: USA
ZIP: 10016
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" FLOPPY DISC
COMPUTER: AT&T - IBM COMPATIBLE
OPERATING SYSTEM: MS-DOS Version 6.2
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/209,603E
FILING DATE: 10-MAR-1994
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP90/02282
FILING DATE: 21-DEC-1990
APPLICATION NUMBER: US 07/862,535
FILING DATE: 19-JUN-1992
ATTORNEY/AGENT INFORMATION:
NAME: ROBINSON, WILLIAM R.
REGISTRATION NUMBER: 27,224
REFERENCE/DOCKET NUMBER: LKR-9217-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 697-3355
TELEFAX: (212) 557-5635
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 210
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
DESCRIPTION: N/A
HYPOTHETICAL: N/A
ANTI-SENSE: N/A
FRAGMENT TYPE: N/A
ORIGINAL SOURCE:
ORGANISM: B. BURGDORFERI
IMMEDIATE SOURCE:
LIBRARY: DSM 5662
FEATURE:
POSITION IN GENOME: N/A
IDENTIFICATION METHOD: amino acid analysis
PUBLICATION INFORMATION: N/A
US-08-209-603E-15

Query Match
Best Local Similarity 33.5%; Score 921; DB 4; Length 210;
Matches 187; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```
Qy    1 MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   17 ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVLEVEARLLSSIDEIAAKA 76

Qy   61 IGKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSEIFTN 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   77 IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSEIFTN 136

Qy  121 KLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  137 KLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 196

Qy  181 ELISPVVAESP 191
         |||||||||||
Db  197 ELISPVVAESP 207
```

RESULT 7
US-08-235-836C-107
Sequence 107, Application US/08235836C
Patent No. 6248562
GENERAL INFORMATION:
APPLICANT: Dunn, John J.
APPLICANT: Luft, Benjamin J.
TITLE OF INVENTION: No. 6248562e1 Chimeric Proteins Comprising
TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor
NUMBER OF SEQUENCES: 144
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brookhaven National Laboratory
STREET:
CITY: Upton
STATE: NY
COUNTRY: USA
ZIP: 11973
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/235,836C
FILING DATE: 29-APR-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/148,191
FILING DATE: 01-11-93
ATTORNEY/AGENT INFORMATION:
NAME: Bogosian, Margaret C.
REGISTRATION NUMBER: 25,324
REFERENCE/DOCKET NUMBER: BNL93-28A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 282-7338
TELEFAX: (516) 282-3729
INFORMATION FOR SEQ ID NO: 107:
SEQUENCE CHARACTERISTICS:
LENGTH: 466 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-235-836C-107

Query Match
Best Local Similarity 33.3%; Score 915; DB 4; Length 466;
Matches 186; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy    4 NNGSKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGK 63
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  276 NNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGK 335

Qy   64 KIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSEIFTNKLK 123
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  336 KIHQNNGLDTEYNHNGSLLAGRYAISTLIKQKLDGLKNEGLKEKIDAAKKCSEIFTNKLK 395

Qy  124 EKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKELT 183
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  396 AKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKELT 455

Qy  184 SPVVAESP 191
         ||||||||
Db  456 SPVVAESP 463
```

RESULT 8
US-08-235-836C-34
Sequence 34, Application US/08235836C

Patent No. 6248562
; GENERAL INFORMATION:
; APPLICANT: Dunn, John J.
; APPLICANT: Luft, Benjamin J.
; TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising
; TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor
; NUMBER OF SEQUENCES: 144
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brookhaven National Laboratory
; STREET:
; CITY: Upton
; STATE: NY
; COUNTRY: USA
; ZIP: 11973
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/235,836C
; FILING DATE: 29-APR-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/148,191
; FILING DATE: 01-11-93
; ATTORNEY/AGENT INFORMATION:
; NAME: Bogosian, Margaret C.
; REGISTRATION NUMBER: 25,324
; REFERENCE/DOCKET NUMBER: BNL93-28A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 282-7338
; TELEFAX: (516) 282-3729
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 212 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-235-836C-34

Query Match   32.3%;   Score 889; DB 4; Length 212;
Best Local Similarity   91.2%; Pred. No. 2.4e-56;
Matches 187; Conservative   1; Mismatches   3; Indels 14; Gaps   2;

```
Qy  195 SNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETLVLSIDELAKKAIG 254
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   20 SNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETLVLSIDELAKKAIG 79

Qy  255 QKIDNNNGLAALNNQNGSLLAGAYAISTLITEKLSKLKNLEELKTEIAKAKKCSEEFTNK 314
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   80 QKIDNNNGLAALNNQNGSLLAGAYAISTLITEKLSKLKNLEELKTEIAKAKKCSEEFTNK 139

Qy  315 LKSGHADLGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGILKAAQVALTNSVKE 374
        |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db  140 LKSGHADLGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGLLKAAQVALTNSVKE 199

Qy  375 LGHRNNSGGDSASTNP--DESAKGP 397
        |:|   |||  || |
Db  200 L---------TSPVVAESPKKP 212
```

RESULT 9
US-08-158-353-4
; Sequence 4, Application US/08158353
; Patent No. 5620862
; GENERAL INFORMATION:
; APPLICANT: Padula, Steven J.
; TITLE OF INVENTION: Methods for Diagnosing Early Lyme
; TITLE OF INVENTION: Disease
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: MA
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/158,353
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Carroll, Alice O.
; REGISTRATION NUMBER: 33,542
; REFERENCE/DOCKET NUMBER: UCT93-05
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 212 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-158-353-4

Query Match   32.2%;   Score 886; DB 1; Length 212;
Best Local Similarity   90.7%; Pred. No. 4e-56;
Matches 186; Conservative   2; Mismatches   3; Indels 14; Gaps   2;

```
Qy  195 SNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETLVLSIDELAKKAIG 254
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   20 NNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETLVLSIDELAKKAIG 79

Qy  255 QKIDNNNGLAALNNQNGSLLAGAYAISTLITEKLSKLKNLEELKTEIAKAKKCSEEFTNK 314
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   80 QKIDNNNGLAALNNQNGSLLAGAYAISTLITEKLSKLKNLEELKTEIAKAKKCSEEFTNK 139

Qy  315 LKSGHADLGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGILKAAQVALTNSVKE 374
        |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db  140 LKSGHADLGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGLLKAAQVALTNSVKE 199

Qy  375 LGHRNNSGGDSASTNP--DESAKGP 397
        |:|   |||  || |
Db  200 L---------TSPVVAESPKKP 212
```

RESULT 10
US-08-235-836C-36
; Sequence 36, Application US/08235836C
; Patent No. 6248562
; GENERAL INFORMATION:
; APPLICANT: Dunn, John J.
; APPLICANT: Luft, Benjamin J.
; TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising
; TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor
; NUMBER OF SEQUENCES: 144
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brookhaven National Laboratory
; STREET:
; CITY: Upton
; STATE: NY
; COUNTRY: USA
; ZIP: 11973
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```
~ OPERATING SYSTEM:  PC-DOS/MS-DOS
~   SOFTWARE:  Patentin Release #1.0, Version #1.25
~ CURRENT APPLICATION DATA:
~   APPLICATION NUMBER:  US/08/235,836C
~   FILING DATE:  29-APR-1994
~   CLASSIFICATION:  435
~ PRIOR APPLICATION DATA:
~   APPLICATION NUMBER:  US 08/148,191
~   FILING DATE:  01-11-93
~ ATTORNEY/AGENT INFORMATION:
~   NAME:  Bogosian, Margaret C.
~   REGISTRATION NUMBER:  25,324
~   REFERENCE/DOCKET NUMBER:  BNL93-28A
~ TELECOMMUNICATION INFORMATION:
~   TELEPHONE:  (516) 282-7338
~   TELEFAX:  (516) 282-3729
~ INFORMATION FOR SEQ ID NO:  36:
~ SEQUENCE CHARACTERISTICS:
~   LENGTH:  207 amino acids
~   TYPE:  amino acid
~   TOPOLOGY:  linear
~   MOLECULE TYPE:  protein
; US-08-235-836C-36

Query Match            32.18%;    Score 882;  DB 4;   Length 207;
Best Local Similarity  99.5%;     Pred. No. 7.4e-56;   Indels   0;   Gaps   0;
Matches 181;   Conservative   1;  Mismatches   0;

Qy   379  NNSGGDSASTNPDESAKGPNLTVISKKIITDSNAFLLAVKEVEALLSSIDELSKAIGKKKIK 438
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    20  NNSGGDSASTNPDESAKGPNLTVISKKIITDSNAFLLAVKEVEALLSSIDELSKAIGKKKIK 79

Qy   439  NDGTLDNEANRNESLIAGAYEISKLITQKLSVLNSEELKKKIKEAKDCSQKFTTKLKDSH 498
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    80  NDGTLDNEANRNESLIAGAYEISKLITQKLSVLNSEELKKKIKEAKDCSEKFTTKLKDSH 139

Qy   499  AELGIQSVQDDNAKRAILKTHGTKDKGAKELEELFKSLESLSKAAQOAALTNSVKELTNPV 558
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   140  AELGIQSVQDDNAKRAILKTHGTKDKGAKELEELFKSLESLSKAAQOAALTNSVKELTNPV 199

Qy   559  VA 560
          ||
Db   200  VA 201
```

```
* RESULT  11
; 9-196-293-11
; Sequence 11, Application US/09196293
; Patent No. 6183755
; GENERAL INFORMATION:
~ APPLICANT:  Fuchs, Renate
~ APPLICANT:  Motz, Manfred
~ APPLICANT:  Soutscheck, Erwin
~ APPLICANT:  Wilske, Bettina
~ APPLICANT:  Preac-Mursic, Vera
~ TITLE OF INVENTION:  Active proteins from Borrelia
~ TITLE OF INVENTION:  burgdorferi
~ FILE REFERENCE:  738.001US2
~ CURRENT APPLICATION NUMBER:  US/09/196,293
~ CURRENT FILING DATE:  1998-11-19
~ EARLIER APPLICATION NUMBER:  US 08/209,603
~ EARLIER FILING DATE:  1994-03-10
~ EARLIER APPLICATION NUMBER:  US 07/862,535
~ EARLIER FILING DATE:  1992-06-19
~ EARLIER APPLICATION NUMBER:  WO PCT/EP90/02282
~ EARLIER FILING DATE:  1990-12-21
~ EARLIER APPLICATION NUMBER:  DE P39 42 728.5
~ EARLIER FILING DATE:  1989-12-22
~ EARLIER APPLICATION NUMBER:  DE P40 18 988.0
~ EARLIER FILING DATE:  1990-06-13
~ NUMBER OF SEQ ID NOS:  16
~ SOFTWARE:  FastSEQ for Windows Version 4.0
```

```
~ SEQ ID NO 11
~   LENGTH:  212
~   TYPE:  PRT
~   ORGANISM:  Borrelia burgdorferi
US-09-196-293-11

Query Match            31.2%;    Score 858;  DB 4;   Length 212;
Best Local Similarity  88.8%;    Pred. No. 3.9e-54;   Indels  14;   Gaps   2;
Matches 182;   Conservative   2;  Mismatches   7;

Qy   195  SNSGKGGDSASTNPADESAKGPNLTEISKKIITDSNAFLVLAVKEVEETLVLSIDELAKKAIG 254
          :||||||  |||||||||||||||||||||||||||||||||||||||||||||||  ||||
Db    20  NNSGKVGILTSTNPADESAKGPNLTEISKKIITDSNAFVLAVKEVEATLVLSIDELAKKAIG 79

Qy   255  QKIDNNNGLAALNNQNGSLLAGAYAISTLLITEKLSKLLKNLEELKTEIAKAKKCSEEFTNK 314
          :|||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    80  QKIDNNNGLAALNNQNGSLLAGAYAISTLLITEKLSKLLKNLEELKTEIAKAKKCSEEFTNK 139

Qy   315  LKSGHADLGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGLLKAAQVALTNSVKE 374
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   140  LKSGHADLGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGLLKAAQVALTNSVKE 199

Qy   375  LGHRNNSGGDSASTNP--DESAKGP 397
          |   | :|:     | | |
Db   200  L--------------TSPVVAESPKKP 212
```

```
* RESULT  12
; US-08-209-603E-11
; Sequence 11, Application US/0820 9603E
; Patent No. 6248538
; GENERAL INFORMATION:
~ APPLICANT:  FUCHS, RENATE
~ APPLICANT:  WILSKE, BETTINA
~ APPLICANT:  PREAC-MURSIC, VERA
~ APPLICANT:  MOTZ, MANFRED
~ APPLICANT:  SOUTSCHEK, ERWIN
~ TITLE OF INVENTION:  IMMUNOLOGICALLY ACTIVE PROTEINS
~ TITLE OF INVENTION:  FROM BORRELIA BURGDORFERI
~ NUMBER OF SEQUENCES:  15
~ CORRESPONDENCE ADDRESS:
~   ADDRESSEE:  BROOKS HAIDT HAFFNER & DELAHUNTY
~   STREET:  99 PARK AVENUE
~   CITY:  NEW YORK
~   STATE:  NY
~   COUNTRY:  USA
~   ZIP:  10016
~ COMPUTER READABLE FORM:
~   MEDIUM TYPE:  3.5" FLOPPY DISC
~   COMPUTER:  AT&T - IBM COMPATIBLE
~   OPERATING SYSTEM:  MS-DOS Version 6.2
~   SOFTWARE:  ASCII
~ CURRENT APPLICATION DATA:
~   APPLICATION NUMBER:  US/08/209,603E
~   FILING DATE:  10-MAR-1994
~   CLASSIFICATION:  436
~ PRIOR APPLICATION DATA:
~   APPLICATION NUMBER:  PCT/EP90/02282
~   FILING DATE:  21-DEC-1990
~   APPLICATION NUMBER:  US 07/862,535
~   FILING DATE:  19-JUN-1992
~ ATTORNEY/AGENT INFORMATION:
~   NAME:  ROBINSON, WILLIAM R.
~   REGISTRATION NUMBER:  27,224
~   REFERENCE/DOCKET NUMBER:  LKR-9217-A
. TELECOMMUNICATION INFORMATION:
~   TELEPHONE:  (212) 697-3355
~   TELEFAX:  (212) 557-5635
~ INFORMATION FOR SEQ ID NO:  11:
~ SEQUENCE CHARACTERISTICS:
~   LENGTH:  212
~   TYPE:  AMINO ACID
```

```
;   TOPOLOGY: LINEAR
;   MOLECULE TYPE: PROTEIN
;   DESCRIPTION:
;   HYPOTHETICAL: N/A
;   ANTI-SENSE: N/A
;   FRAGMENT TYPE: N/A
;   ORIGINAL SOURCE:
;   ORGANISM: B. BURGDORFERI
;   IMMEDIATE SOURCE:
;   LIBRARY: DSM 5662
;   POSITION IN GENOME: N/A
;   FEATURE:
;   IDENTIFICATION METHOD:
;   PUBLICATION INFORMATION: N/A
;   ANTI-SENSE: N/A
;   FRAGMENT TYPE: INTERNAL
;   ORIGINAL SOURCE:
US-08-209-603E-11
```

```
Query Match
Best Local Similarity   88.8%;   Score 858;  DB 4;   Length 212;
Matches 182;  Conservative  2;  Mismatches  7;  Indels  14;  Gaps  2;

Qy  195 SNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETLVLSIDELAKKAIG 254
            :||||                 |||||||||||||||||||||||||||||||||||||
Db   20 NNSGKVGILTSTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETLVLSIDELAKKAIG  79

Qy  255 QKIDNNNGLAALNNQNGSLLAGAYAISTLLTTEKLSKLKNLEELKTEIAKAKKCSEEFTNK 314
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   80 QKIDNNNGLAALNNQNGSLLAGAYAISTLLTTEKLSKLKNLEELKTEIAKAKKCSEEFTNK 139

Qy  315 LKSGHADLGKODATDDHAKAAILKTHATTDKGAKEFKDLFESVEGLLKAAQVALTNSVKE 374
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  140 LKSGHADLGKODATDDHAKAAILKTHATTDKGAKEFKDLFESVEGLLKAAQVALTNSVKE 199

Qy  375 LGHRNNSGGDSASTNP--DESAKGP 397
            |
Db  200 L---------------TSPVVAESPKKP 212
```

```
RESULT 13
US-08-158-353-2
Sequence 2, Application US/08158353
Patent No. 5620862
GENERAL INFORMATION:
;   APPLICANT: Padula, Steven J.
;   TITLE OF INVENTION: Methods for Diagnosing Early Lyme
;   TITLE OF INVENTION: Disease
;   NUMBER OF SEQUENCES: 7
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
;   STREET: Two Militia Drive
;   CITY: Lexington
;   STATE: MA
;   COUNTRY: USA
;   ZIP: 02173
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: Patentin Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/158,353
;   FILING DATE:
;   CLASSIFICATION: 435
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Carroll, Alice O.
;   REGISTRATION NUMBER: 33,542
;   REFERENCE/DOCKET NUMBER: UCT93-05
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 617-861-6240
;   TELEFAX: 617-861-9540
```

```
;   INFORMATION FOR SEQ ID NO: 2:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 212 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: Single
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-08-158-353-2
```

```
Query Match
Best Local Similarity   74.7%;   Score 676;  DB 1;   Length 212;
Matches 145;  Conservative  18;  Pred. No. 3.7e-41;  Mismatches  27;  Indels  4;  Gaps  3;

Qy    1 MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA  60
          ::||||||||||||||||||||||||||||||||:||||||||||:||||:||:||:|| |
Db   17 ISCNNSGKDGNTSANSADESVKGPNLTEISKKITESNAVVLAVKEVEVTLLASIDEVAKKA  76

Qy   61 IGKKIHQNNGLDYEYNHNGSLLAGAYAISTLIKQKLDGIKN-EGLKEKIDAAKKCSEFFT 119
          ||  |  ::|  |   ||||||||||||||:|| ||||:||| :|||||||:|||||:||
Db   77 IGNLIAQ-NGLNAGANQNGSLLAGAYVISTLIAEKLDGLKNSEELKEKIEDAKKCNKAFT 135

Qy  120 NKLKEKHTDLG--KEGVTDADAKEAILKTNGTKTRGAAEELGKLFESVEVLSKAAKEMLAN 177
          :|||  | :||    :||:|||||:||||||||||:||||:|||||||||:::|||||:| |
Db  136 DKLKSSHAELGIANGAATDANAKAAILKTNGTKDKGAQELEKLFESVKNLSKAAQETLNN 195

Qy  178 SVKELTSPVVAESP 191
          |||||||||||:|
Db  196 SVKELTSPVVAENP 209
```

```
RESULT 14
US-08-031-295-2
Sequence 2, Application US/08031295
Patent No. 5530103
GENERAL INFORMATION:
;   APPLICANT: LIVEY, Ian
;   APPLICANT: DORNER, Freidrich
;   TITLE OF INVENTION: METHOD AND COMPOSITION FOR THE
;   TITLE OF INVENTION: PREVENTION OF LYME DISEASE
;   NUMBER OF SEQUENCES: 3
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Foley & Lardner
;   STREET: 3000 K Street, N.W., Suite 500
;   CITY: Washington, D.C.
;   COUNTRY: USA
;   ZIP: 20007-5109
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: Patentin Release #1.0, Version #1.25
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 07/727,245
;   FILING DATE: 11-JUL-1991
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 07/824,161
;   FILING DATE: 22-JAN-1992
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 07/903,580
;   FILING DATE: 25-JUN-1992
;   PRIOR APPLICATION DATA:
;   CLASSIFICATION: 435
;   APPLICATION NUMBER: US/08/031,295
;   FILING DATE: 19930312
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/031,295
;   ATTORNEY/AGENT INFORMATION:
;   NAME: BENT, Stephen A.
;   REGISTRATION NUMBER: 29,768
;   REFERENCE/DOCKET NUMBER: 30472/142 IMMU
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (202)672-5300
;   TELEFAX: (202)672-5399
;   TELEX: 904136
```

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 212 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-031-295-2

Query Match
Best Local Similarity   24.0%;   Score 661;   DB 1;   Length 212;
Matches 142;   Conservative 14;   Mismatches 35;   Indels 2;   Gaps 2;

```
Qy   1 MACNNSGKDGNT-SANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAK 59
         ::||||| |::|||| | |||| |||||||||||||||| :||||||| :|:||||||:|
Db  17 ISCNNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETLVSSIDELATK 76

Qy  60 AIGKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEG-LKEKIDAAKKCSETF 118
         |||||| |||||      |||||||||||||:|  |  |||| | |||||||||||||:|
Db  77 AIGKKIQQNNGLGANADKNGSLLAGAYAISTLITEKLKALKNSGELKAKIEDAKKCSEDF 136

Qy 119 TNKLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANS 178
         | || |  | | ||| || ::|||||||||||||||||| ||| ||| :|||| |
Db 137 TKKLAAGHAQLGIDGATDNDSKEAILKTNGTKTKGAEELVKLSESVASLSKAAQEASANS 196

Qy 179 VKELTSPVVAESP 191
         |||||||||||:|
Db 197 VKELTSPVVAETP 209
```

TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-903-580-2

Query Match
Best Local Similarity   24.0%;   Score 661;   DB 4;   Length 212;
Matches 142;   Conservative 14;   Mismatches 35;   Indels 2;   Gaps 2;

```
Qy   1 MACNNSGKDGNT-SANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAK 59
         ::||||| |::|||| | |||| |||||||||||||||| :||||||| :|:||||||:|
Db  17 ISCNNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETLVSSIDELATK 76

Qy  60 AIGKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEG-LKEKIDAAKKCSETF 118
         |||||| |||||      |||||||||||||:|  |  |||| | |||||||||||||:|
Db  77 AIGKKIQQNNGLGANADKNGSLLAGAYAISTLITEKLKALKNSGELKAKIEDAKKCSEDF 136

Qy 119 TNKLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANS 178
         | || |  | | ||| || ::|||||||||||||||||| ||| ||| :|||| |
Db 137 TKKLAAGHAQLGIDGATDNDSKEAILKTNGTKTKGAEELVKLSESVASLSKAAQEASANS 196

Qy 179 VKELTSPVVAESP 191
         |||||||||||:|
Db 197 VKELTSPVVAETP 209
```

Search completed: March 18, 2002, 09:55:36
Job time: 349 sec

RESULT 15
US-07-903-580-2
Sequence 2, Application US/07903580
Patent No. 6221363
GENERAL INFORMATION:
APPLICANT: LIVEY, Ian
APPLICANT: DORNER, Freidrich
TITLE OF INVENTION: METHOD AND COMPOSITION FOR THE
TITLE OF INVENTION: PREVENTION OF LYME DISEASE
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: Virginia
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/903,580
FILING DATE: 19920625
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/824,161
FILING DATE: 22-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/727,245
FILING DATE: 11-JUL-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/131 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 212 amino acids

THIS PAGE BLANK (USPTO)

OM protein - protein search, using sw model

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

Run on: March 18, 2002, 09:58:33 ; Search time 81.02 seconds
                                               (without alignments)
                                               441.544 Million cell updates/sec

Title: US-09-596-746A-52
Perfect score: 2750
Sequence: 1 MACNNSGKDGNTSANSADES..........KAAQAALTNSVKELTNPVVA 560

Scoring table: BLOSUM62
                    Gapop 10.0 , Gapext 0.5
Searched: 303921 seqs, 63882009 residues

Total number of hits satisfying chosen parameters: 303921

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   Pending_Patents_AA_New:*
               1: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pep:*
               2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep:*
               3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep:*
               4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep:*
               5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep:*
               6: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep1:*
               7: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep:*
               8: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	886	32.2	212	6	US-09-974-992-7	Sequence 7, Appli
2	878	31.9	207	6	US-09-974-992-3	Sequence 3, Appli
3	674.5	24.5	211	6	US-09-974-992-5	Sequence 5, Appli
4	234	8.5	691	5	US-09-748-875-1	Sequence 1, Appli
5	234	8.5	929	5	US-09-748-875-60	Sequence 60, Appli
6	230	8.4	707	5	US-09-748-875-2	Sequence 2, Appli
7	228	8.3	589	5	US-09-748-875-14	Sequence 14, Appli
8	223.5	8.1	690	5	US-09-748-875-61	Sequence 61, Appli
9	223.5	8.1	701	5	US-09-748-875-62	Sequence 62, Appli
10	221	8.0	711	5	US-09-748-875-3	Sequence 3, Appli
11	218.5	7.9	670	5	US-09-748-875-63	Sequence 63, Appli
12	215	7.8	1269	7	US-09-708-427-19881	Sequence 19881, A
13	212	7.7	1881	7	US-10-032-585-7646	Sequence 7646, Ap
14	212	7.7	1881	7	US-10-072-851-15590	Sequence 15590, A
15	209	7.6	1144	6	US-09-708-427-15046	Sequence 15046, A
16	209	7.6	1304	6	US-09-708-427-15044	Sequence 15044, A
17	209	7.6	1313	6	US-09-708-427-15045	Sequence 15045, A
18	201.5	7.3	1690	6	US-09-614-150-10224	Sequence 10224, A
19	201.5	7.3	1690	6	US-09-614-150-10311	Sequence 10311, A
20	197.5	7.2	805	6	US-09-708-427-3730	Sequence 3730, Ap
21	197.5	7.2	841	6	US-09-708-427-3729	Sequence 3729, Ap
22	197.5	7.2	875	6	US-09-815-242-13080	Sequence 13080, A
23	197.5	7.2	875	6	US-10-072-851-13080	Sequence 13080, A
24	197	7.2	6281	6	US-09-815-242-12996	Sequence 12996, A
25	197	7.2	6281	7	US-10-072-851-12996	Sequence 12996, A
26	193.5	7.0	2056	6	US-09-614-150-4824	Sequence 4824, Ap
27	191	6.9	2344	6	US-09-815-242-12713	Sequence 12713, A
28	191	6.9	2344	7	US-10-072-851-12713	Sequence 12713, A
29	189.5	6.9	1014	6	US-10-072-851-19883	Sequence 19883, A
30	189	6.9	1018	6	US-09-708-427-19882	Sequence 19882, A
31	189	6.9	730	6	US-09-708-427-3731	Sequence 3731, Ap
32	188.5	6.9	576775	6	US-08-895-611-2	Sequence 2, Appli
33	188.5	6.9	576775	5	US-08-895-611D-2	Sequence 2, Appli
34	188.5	6.9	576775	6	US-09-895-611D-2	Sequence 2, Appli
35	187	6.8	837	6	US-09-708-427-19883	Sequence 19883, A
36	187	6.8	837	6	US-10-072-851-19882	Sequence 19882, A
37	186	6.8	1489	6	US-09-614-150-6636	Sequence 6636, Ap
38	186	6.8	1786	6	US-09-742-096-3	Sequence 3, Appli
39	185.5	6.7	3158	6	US-09-815-242-5883	Sequence 5883, Ap
40	185.5	6.7	3158	7	US-10-072-851-5883	Sequence 5883, Ap
41	183.5	6.7	2434	6	US-09-815-242-5835	Sequence 5835, Ap
42	183.5	6.7	2434	7	US-10-072-851-5835	Sequence 5835, Ap
43	183.5	6.7	5795	6	US-09-815-242-12610	Sequence 12610, A
44	183.5	6.7	5795	7	US-10-072-851-12610	Sequence 12610, A
45	182.5	6.6	1725	7	US-10-037-182-12	Sequence 12, Appli

ALIGNMENTS

RESULT 1
US-09-974-992-7

Sequence 7, Application US/09974992

; GENERAL INFORMATION:
; APPLICANT: Mathiesen, Marianne J.
; APPLICANT: Theisen, Michael
; APPLICANT: Holm, Arne
; APPLICANT: Ostergaard, Soren
; TITLE OF INVENTION: Novel OspC-derived peptide fragments
; FILE REFERENCE: 459-666P
; CURRENT APPLICATION NUMBER: US/09,974,992
; CURRENT FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: 09/180,089
; PRIOR FILING DATE: 1999-05-13
; PRIOR APPLICATION NUMBER: PCT/DK97/00203
; PRIOR FILING DATE: 1997-05-02
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 7
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Borrelia afzelii
US-09-974-992-7

  Query Match   32.2%; Score 886; DB 6; Length 212;
  Best Local Similarity   90.7%; Pred. No. 7.9e-49;
  Matches 186; Conservative   2; Mismatches   3; Indels 14; Gaps   2;

Qy   195 SNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETLVLSIDELAKKAIG 254
          :|||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db   20 NNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETLVLSIDELAKKAIG 79

Qy   255 QKIDNNNGLAALNNQNGSLLAGAYAISTLITEKLSKLKNLEELKTEIAKAKKCSEEFTNK 314
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   80 QKIDNNNGLAALNNQNGSLLAGAYAISTLITEKLSKLKNLEELKTEIAKAKKCSEEFTNK 139

Qy   315 LKSGHADLGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGLLKAAQVALTNSVKE 374
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   140 LKSGHADLGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGLLKAAQVALTNSVKE 199

Qy   375 LGHRNNSGGDSASTNP--DESAKGP 397
          ||||||||||||||| :|: |
Db   200 L------------TSPVVAESPKKP 212

RESULT 2
US-09-974-992-3

```
;  Sequence 3, Application US/09974992
;  GENERAL INFORMATION:
;  APPLICANT: Mathiesen, Marianne J.
;  APPLICANT: Theisen, Michael
;  APPLICANT: Holm, Arne
;  APPLICANT: Ostergaard, Soren
;  TITLE OF INVENTION: Novel OspC-derived peptide fragments
;  FILE REFERENCE: 459-666P
;  CURRENT APPLICATION NUMBER: US/09/974,992
;  CURRENT FILING DATE: 2001-10-10
;  PRIOR APPLICATION NUMBER: 09/180,089
;  PRIOR FILING DATE: 1999-05-13
;  PRIOR APPLICATION NUMBER: PCT/DK97/00203
;  PRIOR FILING DATE: 1997-05-02
;  NUMBER OF SEQ ID NOS: 40
;  SOFTWARE: PatentIn Ver. 2.1
;
;  SEQ ID NO 3
;  LENGTH: 207
;  TYPE: PRT
;  ORGANISM: Borrelia garinii
US-09-974-992-3

Query Match             31.9%;  Score 878;  DB 6;  Length 207;
Best Local Similarity   98.9%;  Pred. No. 2.4e-48;
Matches   180;  Conservative   2;  Mismatches   0;  Indels   0;  Gaps   0;

Qy    379 NNSGGDSASTNPDESAKGPNLTVISKKITDSNAFLLAVKEVEALLSSIDELSKRAIGKKIK 438
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     20 NNSGGDSASTNPDESAKGPNLTVISKKITDSNAFLLAVKEVEALLSSIDELSKRAIGKKIK 79

Qy    439 NDGTLDNEANRNESLIAGAYEISKLIITQKLSVLNSEELKKKIKEAKDCSQKFTTKLKDSH 498
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db     80 NDGTLDNEANRNESLIAGAYEISKLIITQKLSVLNSEELKEKIKEAKDCSEKFTTKLKDSH 139

Qy    499 AELGIQSVQDDNAKKAILKTHGTRKDKGAKELEELFKSLESLSKAAQAALTNSVKELTNPV 558
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    140 AELGIQSVQDDNAKKAILKTHGTRKDKGAKELEELFKSLESLSKAAQAALTNSVKELTNPV 199

Qy    559 VA 560
          ||
Db    200 VA 201
```

```
RESULT 3
US-09-974-992-5
;  Sequence 5, Application US/09974992
;  GENERAL INFORMATION:
;  APPLICANT: Mathiesen, Marianne J.
;  APPLICANT: Theisen, Michael
;  APPLICANT: Holm, Arne
;  APPLICANT: Ostergaard, Soren
;  TITLE OF INVENTION: Novel OspC-derived peptide fragments
;  FILE REFERENCE: 459-666P
;  CURRENT APPLICATION NUMBER: US/09/974,992
;  CURRENT FILING DATE: 2001-10-10
;  PRIOR APPLICATION NUMBER: 09/180,089
;  PRIOR FILING DATE: 1999-05-13
;  PRIOR APPLICATION NUMBER: PCT/DK97/00203
;  PRIOR FILING DATE: 1997-05-02
;  NUMBER OF SEQ ID NOS: 40
;  SOFTWARE: PatentIn Ver. 2.1
;
;  SEQ ID NO 5
;  LENGTH: 211
;  TYPE: PRT
;  ORGANISM: Borrelia burgdorferi
US-09-974-992-5

Query Match             24.5%;  Score 674.5;  DB 6;  Length 211;
Best Local Similarity   76.7%;  Pred. No. 1.6e-35;
Matches   148;  Conservative   11;  Mismatches   31;  Indels   3;  Gaps   3;
```

```
RESULT 4
US-09-748-875-1
;  Sequence 1, Application US/09748875
;  GENERAL INFORMATION:
;  APPLICANT: BRILES et al.
;  TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS
;  TITLE OF INVENTION: AND STRAINS THEREOF AND USES THEREFOR
;  FILE REFERENCE: 454312-3140
;  CURRENT APPLICATION NUMBER: US/09/748,875
;  CURRENT FILING DATE: 2000-12-26
;  PRIOR APPLICATION NUMBER: 09/298,523
;  PRIOR FILING DATE: 1999-04-23
;  NUMBER OF SEQ ID NOS: 78
;  SOFTWARE: PatentIn Ver. 2.1
;
;  SEQ ID NO 1
;  LENGTH: 691
;  TYPE: PRT
;  ORGANISM: Streptococcus pneumoniae
US-09-748-875-1

Query Match              8.5%;  Score 234;  DB 5;  Length 691;
Best Local Similarity   22.9%;  Pred. No. 4.2e-07;
Matches   144;  Conservative   94;  Mismatches   260;  Indels   132;  Gaps   22;

Qy      1 MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA 60
          ::|||||||||||||||||||||||||||||||||||||||||||||||||||:|| |||
Db     17 ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEL-AKA 75

Qy     61 IGKKIHQNNGLDTEYNHNGSLLAGAYAISLLIKORLDGLK-NEGLKEKIDAAKKCSETFT 119
          |||||||||||||||||||||||||||||||||| | ||| |:| |||||||||||||||
Db     76 IGKKIKNDGSLGDEANHNESLLAGAYTISTLITQKLSKLNGSEGLKEKIAAAKKCSEEFS 135

Qy    120 NKLKEKHTDLGKEGYTDADAKEAILKTNGT-KTKGAEELGKLFESVEVLSKAAKEMLANS 178
          ||:| ||||:|||||||||||||||||||| |:||||||||| |||| |:||||||||||
Db    136 TKLKDNHAQLGIQGYTDENAKKAILKANAAGKDKGVEELKLSGSLESLSKAAKEMLANS 195

Qy    179 VKELTSPVVAESP 191
          ||||||||||||
Db    196 VKELTSPVVVESP 208
```

```
Qy     19 ESVK-GPNLT------------EISKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGKKIH-QNNG 70
          | |  |||||            :||||||||||||||||||||||||||||||||||:| |:| ||
Db     38 EGVRSGNNLTVISSGQDISKKYAD-------------EVESHLESILKDVKNLKKVQHIQNVG 88

Qy     71 LDIE--------------YNHNGSLLAGAYAIS--------TLIKQKLDGLKNEGLKEKIDA 110
          |:|:              :||:||||||||:|         |  | | |:||| ||||| |
Db     89 LITKLSEIKKKYLYDLKVNVLSEAELrSKTKETKEKLTATFEQFKKDTLPTEPEKVVAEA 148

Qy    111 AKKCSETFTNKLKEKHID------------LGKE-GVTDADAKEAILKTNGTRKTKGAEELG 158
          |:|        ||:| |::            |||| ||| ||||||||| ||| | || |
Db    149 QKKVEEAKKAEDQKEKDRRNYPTTITYKTLELEIAESDVEVKKAELELVKVKAKESODEE 208

Qy    159 KLFES-VEVLSK---------AAKEMLANSVKELTSPVVAES----PAMGSNSG 198
          |||:|         |:|        :| | || |||||:||        ||  |
Db    209 KIKQAEAEVESKQAEATRLKIKTDREEAKRKADALKEAVEKNVATSEQDKPKRAKRG 268

Qy    199 KGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETIVLSIDELAKKAIGQ-KI 257
          ||||||||| |::| : :|||||||||||||||:||||||||::|||||||||||| ||
Db    269 VSGELATPDKKENDAKSSD-SSVGEETLPSPSLNMANESQTEHRKDVDEYIKKMLSEIQL 327

Qy    258 DNNNGLAALNNQNGSLLAGAYAISTLITEKLSKIK----NLEEL--------------- 297
          |  |  |  | |::||||||||||||||||  | |    ||||
Db    328 DRRK-----HTQNVNLNIKLSAIKTKYLYELSVLKENSKKEELTSKTKAELTAAFEQFKK 382

Qy    298 ----KTEIAKAKKCSEEF----------------TNKLKSGHADLGKQDATDDHAKAAI 336
          | | |:|:| :| |  :|                ||||:|  |  | |:|| ||||
Db    383 DTLKPEKKVAEAEKKVEEAKKKAKDQKEEDRRNYPTNTYKTLELEIAESDVKKEAELEL 442

Qy    337 LKTHATTDKGAKEFKDLFESVEGLIKAAQVALTNSVKELGHRNNSGGDSASTNPDESAKG 396
          :|:|:|:|:|| |     |:||:|     ||||||||| :|: :||||||||||||||||
Db    443 VKEEANESRNEEKIKQAEKEVES--KKAEATRLEKIKT---DRKAEEEARKAEESEK- 496
```

RESULT 5
US-09-748-875-60
Sequence 60, Application US/09748875
GENERAL INFORMATION:
APPLICANT: BRILES et al.
TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS
TITLE OF INVENTION: AND STRAINS THEREOF AND USES THEREFOR
FILE REFERENCE: 454312-3140
CURRENT APPLICATION NUMBER: US/09/748,875
CURRENT FILING DATE: 2000-12-26
PRIOR APPLICATION NUMBER: 09/298,523
PRIOR FILING DATE: 1999-04-23
NUMBER OF SEQ ID NOS: 78
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 60
LENGTH: 929
TYPE: PRT
ORGANISM: Streptococcus pneumoniae
US-09-748-875-60

```
Query Match        8.5%;
Best Local Similarity   22.9%;   Pred. No. 6.3e-07;
Matches  144;   Conservative  94;   Mismatches  260;   Indels  132;   Gaps  22;

Qy    19  ESVK-GPNLT-----EISKKITDSNAVLLAVKEVEALLSSIDELAAKAIGKKIH-QNNG   70
Db    38  EGVRSGNNLTVTSSGQDISKKYAD------EVESHLESILKDVKKNLKKVQHTQNVG     88

Qy    71  LDTE-------YNHNGSLLAGAYAIS------TLIKQKLDGLKNEGLKEKIDA        110
Db    89  LITKLSEIKKKYLYDLKVNVLSEAELTSKTKEKLTATFEQFKKDTLPTEPEKKVAEA    148

Qy   111  AKKCSSETFTNKLKEKHTD------LGKE-GVTDADAKEAILKTNGTKTKGAEELG     158
Db   149  QKKVEEAKKAEDQKEKDRRNYPTITYKTLELEIAESDVEVKKAELELVKVKAKESQDEE   208

Qy   159  KLFES-VEVLSK------AAKEMLANSVKELTSPVVAES-----PAMGSNSG         198
Db   209  KIKQAEAEVESKQAEATRLKKIKTDREEAKRKADAKLKEAVEKNVATSEQDKPKRRAKRG  268

Qy   199  KGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETLVLSIDELAKKAIGQ-KI  257
Db   269  VSGELATPDKKENDAKSSD-SSVGEETLPSPSLNMANESQTEHRKDVDEYIKKMLSEIQL  327

Qy   258  DNNNGLAALNNQNGSLLAGAYAISTLITEKLSKLK---NLEEL------            297
Db   328  DRRK----HTQNVLNIKLSAIKTKYLYELSVLKENSKKEELTSKTKAELTAAFEQFKK   382

Qy   298  -----KTEIARAKKCSEEF------TNKLKSGHADLGKQDATDDHAKAAI           336
Db   383  DTLKPEKKVAEAEKKVEEAKKKAKDQKEEDRRNYPTNTYKTLELEIAESDVKVKEAELEL  442

Qy   337  LKTHATTDKGAKEFKDLFESVEGLLKAAQVALTNSVKELGHRNNSGGDSASTNPDESAKG  396
Db   443  VKEEANESRNEEKIKQAKEKVES--KKAEATRLKEKIKT---DRKKAEEEAKRKAEESEK- 496

Qy   397  PNLTVISKKITDSNAFLLAVKEVEALLSSIDELSKAIGKKIKNDGTLDNEANRNESLIAG  456
Db   497  --KAAEAKQKVDAEEYAL-----EAKIAELEYEVQRLEKELKEIDESDSEDYLKEGLRAP  549

Qy   457  AYEISKLITQKLSVLNSEELKKKIKEAKDCSQKFTTKLKDSHAELGIQSVQDDNAKKAIL  516
Db   550  LQ--SKLDTKKAKLSKLEELSDKIDELDAEIAKLEVQLKDAEGNNNVEAYFKEGLEKTTA  607

Qy   517  KTHGTKDKGAKELEELFKSLESLSKRAAQAA                             546
Db   608  EKKAELEKAEADLKKAVDEPETPAPAPQOPA                             637
```

RESULT 6
US-09-748-875-2
Sequence 2, Application US/09748875
GENERAL INFORMATION:
APPLICANT: BRILES et al.
TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS
TITLE OF INVENTION: AND STRAINS THEREOF AND USES THEREFOR
FILE REFERENCE: 454312-3140
CURRENT APPLICATION NUMBER: US/09/748,875
CURRENT FILING DATE: 2000-12-26
PRIOR APPLICATION NUMBER: 09/298,523
PRIOR FILING DATE: 1999-04-23
NUMBER OF SEQ ID NOS: 78
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 2
LENGTH: 707
TYPE: PRT
ORGANISM: Streptococcus pneumoniae
US-09-748-875-2

```
Query Match        8.4%;
Best Local Similarity   22.2%;   Pred. No. 7.7e-07;
Matches  139;   Conservative  100;   Mismatches  264;   Indels  122;   Gaps  21;

Qy    19  ESVK-GPNLT-----EISKKITDSNAVLLAVKEVEALLSSIDELAAKAIGKKIH-QNNG   70
Db    38  EGVRSGNNLTVTSSGQDISKKYAD------EVESHLESILKDVKKNLKKVQHTQNVG     88

Qy    71  LDTE-------YNHNGSLLAGAYAIS------TLIKQKLDGLKNEGLKEKIDA        110
Db    89  LITKLSEIKKKYLYDLKVNVLSEAELTSKTKEKLTATFEQFKKDTLPTEPEKKVAEA    148

Qy   111  AKKCSSETFTNKLKEKHTD------LGKE-GVTDADAKEAILKTNGTKTKGAEELG     158
Db   149  QKKVEEAKKAEDQKEKDRRNYPTITYKTLELEIAESDVEVKKAELELVKVKAKESQDEE   208

Qy   159  KLFES-VEVLSK------AAKEMLANSVKELTSPVVAES-----PAMGSNSG         198
Db   209  KIKQAEAEVESKQAEATRLKKIKTDREEAKRKADAKLKEAVEKNVATSEQDKPKRRAKRG  268

Qy   199  KGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETLVLSIDELAKKAIGQ--  255
Db   269  VSGELATPDKKENDAKSSD-SSVGEETLPSPSLNMANESQTEHRKDVDEYIKKMLSEIQL  327

Qy   256  ------KIDNNNGLAAL------                                     290
Db   328  DGRKHTPNVNLNIKLSAIKTKYLYELSVLKENSKKEELTSKTKAELTAAFEQFKKDTLKP  387

Qy   291  LKNLEELKTEIARAKKCSEEF------TNKLKSGHADLGKQDATDDHAKAAILKTHA   341
Db   388  EKKVAEAEKKVEEAKKKAKDQKEEDRRNYPTNTYKTLELEIAESDVKVKEAELELVKEEA  447

Qy   342  TTDKGAKEFKDLFESVEGLLKAAQVALTNSVKELGHRNNSGGDSASTNPDESAKGPNLTV  401
Db   448  NESRNEEKIKQAKEKVES--KKAEATRLEKIKT---DRKKAEEEAKRKAEESEK---KAA  499

Qy   402  ISKKITDSNAFLLAVKEVEALLSSIDELSKAIGKKIKNDGTLDNEANRNESLIAGAYEIS  461
Db   500  EAKQKVDAEEYAL-----EAKIAELEYEVQRLEKELKEIDESDSEDYLKEGLRAPLQ--S  552

Qy   462  KLITQKLSVLNSEELKKKIKEAKDCSQKFTTKLKDSHAELGIQSVQDDNAKKAILKTHGT  521
```

RESULT 7
US-09-748-875-14
; Sequence 14, Application US/09748875
; GENERAL INFORMATION:
; APPLICANT: BRILES et al.
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS
; TITLE OF INVENTION: AND STRAINS THEREOF AND USES THEREFOR
; FILE REFERENCE: 454312-3140
; CURRENT APPLICATION NUMBER: US-09/748,875
; CURRENT FILING DATE: 2000-12-26
; PRIOR APPLICATION NUMBER: 09/298,523
; PRIOR FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 589
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae

Query Match   8.3%; Score 228; DB 5; Length 589;
Best Local Similarity   23.6%; Pred. No. 8.1e-07;
Matches 148; Conservative 89; Mismatches 247; Indels 142; Gaps 24;

```
Qy     19 ESVK-GPNLT--------EISKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGKKIH-QNNG  70
             ::  |||         |||||||||||||||||||||||||||||||||||| |||
Db      1 EGVRSGNNLIVTSSGQDISKKYAD--------EVESHLESILKDVKNLKKVQHTQNVG  51

Qy     71 LDTE----------------YNHNGSLLAGAYAIS----------TLIKQKLDGLKNEGLKEKIDA 110
           |:|:                  ||||                  | |:
Db     52 LITKLSEIRKKYLYDLKVNVLSEAELTSKTKETKEKLTATFEQFKKDTLPTEPEKKVAEA 111

Qy    111 AKKCSETFTNKLKEKHTD----------IGKE-GVTDADAKEAILKTNGTKTKGAEELG 158
           ||                        |||| |
Db    112 QKKVEEAKKKAEDQKEKDRRNYPITIYKTLELEIAESDVEVKRAELVKVKAKESQDEE 171

Qy    159 KLFES-VEVLSK---------------AAKEMLANSVKELTSPVVAES-----PAMGSNSG 198
           ||  |   |||                |:| | |         |||       |
Db    172 KIKQAEAEVESKQAEEATRLRKKIKIDREEAKRKADAKLKEAVEKNVATSEQDKPKRRARKG 231

Qy    199 KGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETLVLSIDELAKRAIGQ-KI 257
           |||||||||||||||||||||||||||||||||||||||||||||| |  :   |
Db    232 VSGELATPDKKENDAKSSD--SSVGEETLPSPSLNMANESQTEHRKDVDEYIKKMLSEIQL 290

Qy    258 DNNNGLAALNNQNGSLLAGAYAISTLIITEKLSKLK---NLEEL--------------- 297
           ||||          | |      |:| | |          ||||
Db    291 DRRK-----HTQNVLNIKLSAIKTKLYLELSVLKENSKKEELTSKITKAELTAAFEQFKK 345

Qy    298 -----KTEIAKAKKCSEEF----------------------------------------- 298
                :||||||||||| |
Db    346 DTLKPEKKVAEAKKRAKDQKEEDRRNYPTNTYKTLELEAESDVKVKEAELEL 405

Qy    337 LKTHATTDKGAKEFKDLFESVEGLLKAAQVALTNSVKELGHRNNSGGDSASTNPDESAKG 396
           | | ||| |        |:   |         |  |          ||
Db    406 VKEEANESRNEEKIKQAREKVES--KRAEATRLEKIKT--DRKKAEEEAKRKAEESEK- 459

Qy    397 PNLTVISKKITDSNAFLLAVKEVEALLSSIDELSKAIGKKIKNDGTLDNEANRNESLIAG 456
           | | :| ||                 |   |          |:
Db    460 --KAAEAKQVDAEEYAL------EAKIAELEVEVQRLEKELKEIDESDSEDYLKEGLRAP 512

Qy    457 AVEISKLITQKLSVLNSEELKKKIKEARDCSQKFTTKLKDSHAELGIQSVQDDNAKKAIL 516
           | |:|    |   | |   |   |    ||::| | ||    |   | :    | | |
Db    513 LQ--SKLDTKKAKLSKLEELSDKIDELDAEIAKLEVQLKDA----------EGNNNVEAYF 561

Qy    517 KTHGTKDKGAKELELFKSLESLSKA 542
           |  | ||| ||| || ||||| |||
Db    562 K-EGLEKITAEKKAELEKAEAADLKKA 586

Qy    553 KLDDTKKAKLSKLEELSDKIDELDAEIAKLEVQLKDAEGNNNVEAYFKEGLEKITAEKKAE 612
           ||:|  || |||| || | ||  |||| ||       ||    |  : :| ||
Db    613 LEKAEADLKKAVDEPETPAPQPA 637

Qy    522 KDKGAKELELFKSLESLSKAQQAA 546
           || |||| ||| |||||| |
Db    553 ...
```

RESULT 8
US-09-748-875-61
; Sequence 61, Application US/09748875
; GENERAL INFORMATION:
; APPLICANT: BRILES et al.
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS
; TITLE OF INVENTION: AND STRAINS THEREOF AND USES THEREFOR
; FILE REFERENCE: 454312-3140
; CURRENT APPLICATION NUMBER: US-09/748,875
; CURRENT FILING DATE: 2000-12-26
; PRIOR APPLICATION NUMBER: 09/298,523
; PRIOR FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 61
; LENGTH: 690
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae

Query Match   8.3%; Score 223.5; DB 5; Length 690;
Best Local Similarity   22.7%; Pred. No. 1.9e-06;
Matches 143; Conservative 92; Mismatches 263; Indels 131; Gaps 22;

```
Qy     19 ESVK-GPNLT---------EISKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGKKIH-QNNG 70
             ::  |||         |||||||||||||||||||||||||||||||||||| |||
Db     38 EGVRSGNNLIVTSSGQDISKKYAD--------EVESHLESILKDVKNLKKVQHTQNVG 88

Qy     71 LDTE----------------YNHNGSLLAGAYAIS----------TLIKQKLDGLKNEGLKEKIDA 110
           |:|:                  ||||                  | |:
Db     89 LITKLSEIRKKYLYDLKVNVLSEAELTSKTKETKEKLTATFEQFKKDTLPTEPEKKVAEA 148

Qy    111 AKKCSETFTNKLKEKHTD----------IGKE-GVTDADAKEAILKTNGTKTKGAEELG 158
           ||                        |||| |
Db    149 QKKVEEAKKKAEDQKEKDRRNYPITIYKTLELEIAESDVEVKRAELVKVKAKESQDEE 208

Qy    159 KLFES-VEVLSK---------------AAKEMLANSVKELTSPVVAES-----PAMGSNSG 198
           ||  |   |||                |:| | |         |||       |
Db    209 KIKQAEAEVESKQAEEATRLRKKIKIDREEAKRKADAKLKEAVEKNVATSEQDKPKRRARKG 268

Qy    199 KGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETLVLSIDELAKRAIGQ-KI 257
           |||||||||||||||||||||||||||||||||||||||||||||| |  :   |
Db    269 VSGELATPDKKENDAKSSD--SSVGEETLPSPSLNMANESQTEHRKDVDEYIKKMLSEIQL 327

Qy    258 DNNNGLAALNNQNGSLLAGAYAISTLIITEKLSKLK---NLEEL--------------- 297
           ||||          | |      |:| | |          ||||
Db    328 DRRK-----HTQNVLNIKLSAIKTKLYLELSVLKENSKKEELTSKITKAELTAAFEQFKK 382

Qy    298 -----KTEIAKAKKCSEEFTNKLKSGHAD------------------------------ 298
                :||||||||||| |      |
Db    383 DTLKPEKKVAEAKKRAKDQKEEDRRNYPTNTYKTLELEAESDVKVKEAELELV 442

Qy    338 KTHATTDKGAKEFKDLFESVEGLLKAAQVALINSVKELGHRNNSGGDSASTNPDESAKGP 397
           | | ||| |        |:   |         |  |          ||
Db    443 KEEANESRNEEKIKQAREKVES--KRAEATRLEKIKT--DRKKAEEEAKRKAEESEK-- 495

Qy    398 NLTVISKKITDSNAFLLAVKEVEALLSSIDELSKAIGKKIKNDGTLDNEANRNESLIAGA 457
           | | :| ||                 |   |          |:
Db    496 -KAAEAKQVDAEEYAL------EAKIAELEVEVQRLEKELKEIDESDSEDYLKEGLRAPL 549

Qy    458 YEISKLITQKLSVLNSEELKKKIKEARDCSQKFTTKLKDSHAELGIQSVQDDNAKKAILK 517
           |:|    |   | |   |   |    ||::| | ||    |   | :    | | |
Db    550 Q--SKLDTKKAKLSKLEELSDKIDELDAEIAKLEVQLKDA----------EGNNNVEAYF 607

Qy    518 THGTKDKGAKELELFKSLESLSKAQQAA 546
           |  | ||| ||| || ||||| |||
Db    608 KKAELEKAEADLKKAVDEPETPAPQPA 636
```

```
RESULT 9
US-09-748-875-62
  Sequence 62, Application US/09748875
; GENERAL INFORMATION:
;   APPLICANT: BRILES et al.
;   TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS
;   TITLE OF INVENTION: AND STRAINS THEREOF AND USES THEREFOR
;   FILE REFERENCE: 454312-3140
;   CURRENT APPLICATION NUMBER: US/09/748,875
;   CURRENT FILING DATE:
;   PRIOR APPLICATION NUMBER: 09/298,523
;   PRIOR FILING DATE: 1999-04-23
;   NUMBER OF SEQ ID NOS: 78
;   SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 62
;   LENGTH: 701
;   TYPE: PRT
;   ORGANISM: Streptococcus pneumoniae
US-09-748-875-62

  Query Match             8.1%;  Score 223.5;  DB 5;  Length 701;
  Best Local Similarity   22.7%;  Pred. No. 2e-06;  Mismatches 263;  Indels 131;  Gaps 22;
  Matches  143;  Conservative   92;

Qy   19 ESVK-GPNLT------------EISKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGKKIH-QNNG  70
           : | |:|||                |||||||||||||||||||||||||||||||||||  ||||
Db   38 EGVRSGNNLTVTSSGQDISKKYAD------------EVESHLESILKDVKRNLKKVQHTQNVG  88

Qy   71 LDTE------------------YNHNGSLLAGAYAIS------------TLIKQKLDGLKNEGLKEKIDA 110
          :                     :  ||: ||| ||:|                  : | ||:  |  :|:| |
Db   89 LITKLSEIKKKYLYDLKVNVLSEAELTSKTKETKEKLTATFEQFKKDTLPTEPEKKVAEA 148

Qy  111 AKKCSEFFTNRLKEKHTD------------LGKE-GVTDADAKEAILKTNGTKTKGAEELG 158
         |:|   :|  ||||:||                ||:| |||:||||:||:|||||||
Db  149 QKKVEEAKKAEDQKEKDRRNYPTITYKTLELEIAESDVEVKKAELELVKVKAKESQDEE 208

Qy  159 KLFES-VEVLSK------------AAKEMLANSVKELTSPVVAES-----PAMGSNSG 198
          :||:  :|:|              |||:|||||||          : |  ||  ||
Db  209 KIKQAEAEVESKQAEATRLKKIKTDREEAAKRKADAKLKEAVEKNVATSEQOKPKRAKRG 268

Qy  199 KGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETTLVLSIDELAKKAIGQ-KI 257
          ||||||||||||||||||||||||||||||||||||||||  :|: ||||||||||| ||
Db  269 VSGELATPDKKENDAKSSD-SSVGEETLPSPSLNMANESQTEHRKDVDEYIKKMLSEIQL 327

Qy  258 DNNGLAALNNQNGSLLAGAYAISTLITEKLSKLK----NLEEL------- 297
          |  |  |                            || ||||
Db  328 DRRK-----HTQNVLNIKLSAIKTKYLYELSVLKENSKKEELTSKTKAELTAAFEQFKK 382

Qy  298 -----KTEIAKAKCSEEFTNKLKSGHAD-----------LGKQDATDDHAKAAIL 337
              |||| | |  |   :|| |                   ||  | :|
Db  383 DTLKPEKVYAEAKKVEEAKKKAKDQKEEDRRNYPTNTYKTLELEAESDVKVEAELELV 442

Qy  338 KTHATTDKGAKEFKDLFESVEGLLKAAQVALTNSVKELGHRNSGGDSASTNPDESAKGP 397
          ||  | ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  443 KEEANESRNEEKIKQAKEKVES---KKAEATRLEKIKT---DRKKAEEEAARKKAEESEK-- 495

Qy  398 NLTVISKKITDSNAFLLAVKEVEALLSSIDELSKAIGKKIKNDGTLDNEANRNESLIAGA 457
          ||:  |   :|  | |:|  |    :  :| | |    |:|:
Db  496 -KAAEAKQKVDAEEYAL----EAKIAELEYEVQRLEKELKEIDESDSEDYLKEGLRAPL 549

Qy  458 YEISKLITQKLSVLNSEELKKKIKEAKDCSQKFTTKLKDSHAELGIQSVQDDNAKKAILK 517
          : |:|| ||||:||||||:  | ||   |  || ||:|||||||||| | :||||| ||:
Db  550 Q--SKLDTKRAKLSKLEELSDKIDELDAEIAKLEVQLKDAEGNNNVEAYFKEGLEKTTAE 607

Qy  518 THGTKDKGAKELEELFKSLESLSKAAQOA 546
          |  ||||||||||||||||||||||| |
Db  608 KKAELEKAEADLKKAVDEPETPAPAPQPA 636

RESULT 10
```

```
US-09-748-875-3
  Sequence 3, Application US/09748875
; GENERAL INFORMATION:
;   APPLICANT: BRILES et al.
;   TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS
;   TITLE OF INVENTION: AND STRAINS THEREOF AND USES THEREFOR
;   FILE REFERENCE: 454312-3140
;   CURRENT APPLICATION NUMBER: US/09/748,875
;   CURRENT FILING DATE: 2000-12-26
;   PRIOR APPLICATION NUMBER: 09/298,523
;   PRIOR FILING DATE: 1999-04-23
;   NUMBER OF SEQ ID NOS: 78
;   SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 3
;   LENGTH: 711
;   TYPE: PRT
;   ORGANISM: Streptococcus pneumoniae
US-09-748-875-3

  Query Match             8.0%;  Score 221;  DB 5;  Length 711;
  Best Local Similarity   20.5%;  Pred. No. 2.9e-06;  Mismatches 258;  Indels 144;  Gaps 20;
  Matches  131;  Conservative  106;

Qy   12 TSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAIG-----KKIH  66
          ||:| |||                                                  |||
Db   48 TSYNKANES----------QTEHRK-----AAKQVDEDIKKMLSEIQEYIKKMLSEIQLDKRRH  96

Qy   67 QNN-----GLDTEYNHNGSLL---AGAYAISTLIKQKLDGLKNEGLKEKIDAAKKGS 115
          ||:     ||:|                :|| | |:| | |||:|||| |:| ||
Db   97 TQNVNLNRKLSAIQTKYLYELRVLKEKSKKEELTSKTKKELDAAFEKFKKEEPELTKLA 156

Qy  116 ET------------FTNKLKEKHTDLGK-------------EGVTDAD 138
          |:            |||||||||||||                |||||
Db  157 EAKQKAKAQKEEDFRNYPTNTYKTLELEIAEFDVKVKEAELELVKEEAKPRNEEKIKQAK 216

Qy  139 AKEAILKTNGT----------KTKGAEELGKLFESVEVLSKAAKEMLANSVK------- 180
          |:| ||:|              |||| | | | | | ::|  |:|:||||:|
Db  217 AKVESKKAEATRLEEIKTERKAEEEAARKKAEESEKKAAEAKQKVDTKEQGKPKRAKRG 276

Qy  181 ---ELTSPVVAESPAMGSNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKE 237
             |||  :         |  ||||||||||||||||||||||||||||||||||||||
Db  277 VSGELATPDKKENDAKSSDSSVG-----------EETLPSPSLNMANESQTEHR-----KD 322

Qy  238 VETLV--------LSIDELAKKAIGQKIDNNGLAAL-------NNQNGSLLAGAYAIS 281
          ||                 ||||||| :| | ||         |||| |
Db  323 VDEYIKKMLSEIQLDRRKHTQNVNLNIKLSAIKTKYLYELSVLKENSKKEELTSKTKAEL 382

Qy  282 TLITEKLSK--------LKNLEELKTEIAKAKKCSEEF--------TNKLKSGHADLGKDA 327
          ||                      ||||| | |  ::||        |||||
Db  383 TAAFEQFKDTLKPEKVYAEAKKVEEAKKKAKDQKEEDRRNYPTNTYKTLELEIAESDV 442

Qy  328 TDDHAKAAILKTHATTDKGAKEFKDLFESVEGLLKAAQVALTNSVKELGHRNSGGDSAS 387
          |::|  | |||  | | ||||||||||||||||||||||||||||||||||||||||||
Db  443 KVKAELELVKEEANESRNEEKIKQAKEKVES--KKAEATRLEKIKT---DRKKAEEEAK 497

Qy  388 TNPDESAKGPNLTVISKKITDSNAFLLAVKEVEALLSSIDELSKAIGKKIKNDGTLDNEA 447
          |||| ||:||||||||||||||||||||||||||||||||||||||||||||| |:|:|
Db  498 RKAEESEK---KAAEAKQKVDAEEYAL-----EAKIAELEYEVQRLEKELKEIDESDSED 549

Qy  448 NRNESLIAGAYEISKLITQKLSVLNSEELKKKIKEAKDCSQKFTTKLKDSHAELGIQSVQ 507
          |:|||| |||: |:|| ||||:||||||:  | ||   |  || ||:|||||||||| |
Db  550 YLKEGLRAPLQ--SKLDTKRAKLSKLEELSDKIDELDAEIAKLEVQLKDAEGNNNVEAYF 607

Qy  508 DDNAKKAILKTHGTKDKGAKELEELFKSLESLSKAAQOA 546
          :|||| |||  |  ||||||||||||||||||||||| |
Db  608 KEGLEKTTAEKKAELEKAEADLKKAVDEPETPAPAPQPA 646

RESULT 11
US-09-748-875-63
  Sequence 63, Application US/09748875
; GENERAL INFORMATION:
```

Left column

```
APPLICANT: BRILES et al.
TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS
TITLE OF INVENTION: AND STRAINS THEREOF AND USES THEREFOR
FILE REFERENCE: 454312-3140
CURRENT APPLICATION NUMBER: US/09/748,875
CURRENT FILING DATE: 2000-12-26
PRIOR APPLICATION NUMBER: 09/298,523
PRIOR FILING DATE: 1999-04-23
SOFTWARE: Patentin Ver. 2.1
NUMBER OF SEQ ID NOS: 78
SEQ ID NO 63
  LENGTH: 670
  TYPE: PRT
  ORGANISM: Streptococcus pneumoniae
US-09-748-875-63

Query Match
Best Local Similarity   7.98;  Score 218.5;  DB 5;  Length 670;
Matches 142;  Conservative 81;  Pred. No. 3.8e-06;  Mismatches 248;  Indels 151;  Gaps 23;

Qy   12 TSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAIG---KKIH 66
            ||| |:|| ||| ||||||||||||||||||||||||||||||||| |   |:||
Db   48 TSYNKANES-----QTEHRK-----AAKQVDEDIKKMLSEIQEYIKKMLSEIQLDKRKH 96

Qy   67 QNNGLDTEYNHNGSLIAGAYAISTLIKQKLDGLKNEGLKEKIDA------AKKCSE---T 117
            || |:| ||||||| ||| |:|||||| ||:|||| |||:         |||||   |
Db   97 TON----VNLNRKL---SAIQTKYLYELVLKEKSKKEELTSNYPTNIYKTLELEIAE 147

Qy  118 FTNKLKEKHIDLGK-------EGVTDADAKEAILKTNGT--------KTKGAELGKLF 161
            |||| || |               |||| | ||| |  |         |:|||| |||
Db  148 FDVKVKEAELEVKEEAKPRNEEKIKQAKAKVESKKAEATRLEEIKTERKKAEEEAKRKA 207

Qy  162 ESVEVLSKAAKEMLANSVK----ELTSPVVAESPAMGSNSGKGGD---SAS 205
            ||||||||||||||||||    ||:||||||||||||| |||| |   |||
Db  208 EESEKKAAEAKQKVDITKEQGKPKIRAKRGVSGELATPDKKENDAKSSDSSVGEETLPSPS 267

Qy  206 TNPADESAKGPNLTEISKKITDSNAFVLAVKEVETLVLSIDELAKKAIGQ-KIDNNNGLA 264
            |:||||||| ||:||| |||| |:||  | ||||||| | |:|||| | | |||||||||
Db  268 LNMANESQ-----TEHRK------------------DVDEYIKKMLSEIQLDRRK--- 299

Qy  265 ALNNONGSLLAGAYAISTLITEKLSKLK---NLEEL------KT 299
            ||:| |||||||||||||||||||||    ||:||      ||
Db  300 --HTQNVLNIKLSAIKTKYLYELSVLKENSKKEELTSKTKAELTAAFEQFKKDTLKPEK 357

Qy  300 EIAKAKKCSEEFTNKLKSGHAD-------------LGKQDATDDHAKAAILKTHATTD 344
            ||||||||||||||||| |               |||||||||||||||||| | |
Db  358 KVAEAEKKVEEAKKKAKDQKEEDRRNYPTNTYKTLELEAESDVKDKAELELVKEEANES 417

Qy  345 KGAKEFKDLFESVEGLLKAAQVALTNSVKELGHRNSGGDSASTNPDESAKGPNLTVISK 404
            ||||||||||||||||||||||:|||||||||| | ||:||||||||||||||||||||
Db  418 RNEEKIKQAKEKVES--KKAEATRLEKIKT--DDRKAEEEAKRKAEESEK---KAAEAK 470

Qy  405 KITDSNAFLIAVKEVEALLSSIDELSKAIGKKIKNDGTLDNEANRNESLIAGAYEISKLI 464
            ||||||||| || ||||||||||||||||||| ||||| |:|||||||||||||||||
Db  471 QKVDAEEYAL----EAKIAELEYEVQRLEKELKEIDESDSEDYLKEGLRAPLQ--SKID 523

Qy  465 TQKLSVLNSEELKKKIKEAKDCSQKFTFKLKDSHAELGIQSVQDDNAKKAILKTHGTKDK 524
            | |||||||| || | || | ||| ||||||| | |||||||||||| ||||||| |||
Db  524 TKKAKLSKLEELSDKIDELDAEIAKLEVQLKDAEGNNNVEAYFKEGLEKTAEKKAELEK 583

Qy  525 GAKELEELFKSLESLSKRAAQA 546
            |:|||||||||||||| |:| |
Db  584 AEADLKKAVDEPETPAPAPQPA 605
```

Right column

```
US-09-596-746a-52.rapn

FILE REFERENCE: 2750-1243P
CURRENT APPLICATION NUMBER: US/09/708,427
CURRENT FILING DATE: 2000-11-09
NUMBER OF SEQ ID NOS: 85364
SOFTWARE: Patentin version 3.1
SEQ ID NO 19881
  LENGTH: 1269
  TYPE: PRT
  ORGANISM: Arabidopsis thaliana
  FEATURE:
  NAME/KEY: misc_feature
  LOCATION: 1..1269
  OTHER INFORMATION: Xaa is any amino acid
  FEATURE:
  NAME/KEY: misc_feature
  LOCATION: 1..1269
  OTHER INFORMATION: Ceres Seq. ID 1836845
US-09-708-427-19881

Query Match
Best Local Similarity   7.8%;  Score 215;  DB 6;  Length 1269;
Matches 155;  Conservative 118;  Pred. No. 1.5e-05;  Mismatches 255;  Indels 172;  Gaps 33;

Qy   13 SANSAD-ESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGKKIHQNNGL 71
Db  224 SASHADSESQKALEFSELLKSTKES--------AKEMEKMASLQQ-EIKELNEKMSENEKV 276

Qy   72 DTEYNHNGSLIAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSEFFTNKLKEKHTDLG- 130
Db  277 EAALKSSAGELAAVQEELALSKSRL--LETE---QKVSSTEALLIDELTQELEQKKASESR 331

Qy  131 -KEGVT---DADAKEAILKT------NGTKTKGAELGKLFESVEVLSKAAKEMLANSVKE 181
Db  332 FKEELSVLQDLDAQTKGLQAKLSEQEGINSKLAEEL-KEKELLESLSKDQEEKL-RIANE 389

Qy  182 LTSPVVAESPAMGSNSGKGGDSAST-----NPADESAKG-------PNLTE 220
Db  390 KLAEVILKEKEALEANVAEVTSNVATVTEVCNELEEKLKTSDENFSKTDALLSQALSNNSE 449

Qy  221 ISKKITD------SNAFVLAVK-----EVETLVLSIDELAKKAIGQKIDNNNGLAALNNQ 269
Db  450 LEQKLKSLEELHSEAGSAAAAATQKNLELEDVVRSSSQQAAFEAKSQIKELETKFTAAEQK 509

Qy  270 NGSLLAGAYAISTLITEKLSKLKNLEELKTEIAKAKKCSEEFTN------KLKSGHA 320
Db  510 NAELEQQLNLLQLKSSDAERELKELSERKSSELQTAIEVAEEEKKQATTQMQEYKQKASEL 569

Qy  321 DL------GKQDATDDHAKAAILK--THA----TTDKGAKEFKDL------FESVEGLL 361
Db  570 ELSLTQSSARNSELEEDLRIALQKGAEHEDRANTTHQRSIELEGLCQSSQSKHEDAEGRL 629

Qy  362 KAAQVALTNS----VKELGHRNNS------------GGDS---- 385
Db  630 KDLELLLQTEKYRIQELEEQVSSLEKKHGETEADSKGYLGQVAELQSTLEAFQVKSSSLE 689

Qy  386 -----ASTNPDESAKGPNLTVISKKITDS--NAFLLAVKEVEALLSSI-DELSKAIGK-- 435
Db  690 AALNIATENEKELTENILNAVTSEKKLEATVDEYSVKISESENLLSSIRNELNVTQGKLE 749

Qy  436 KIKND---GTLDNEA----NRNESLIAGAYEISKLIT------------QKLSVLNSEELK 477
Db  750 SIENDLKAAGLQESEVMEKLKSAEESLEQKGREIDEATTKRMELEALHQSLSIDSEHRLQ 809

Qy  478 KKIK------EAKDCSQKFTT---KLKDSHAELGIQSVQDDNAKKAILKTHG------ 520
Db  810 KAMEEFTSRDSEAASLTEKLRDLEGKIKSYEEQLAEASGKSSSLIKEKLEQTLGRLAAAES 869

Qy  521 TKDKGAKELEEL-FKSLESLSKAAQAAILTNS-----VKEL 554
Db  870 VNEKLKQEFDAQQEKSLQSSSESELLAETNNQLKIKIQEL 909
```

RESULT 13

```
US-10-032-585-7646
Sequence 7646, Application US/10032585
GENERAL INFORMATION:
;  APPLICANT: Terry, Roemer D.
;  APPLICANT: Bo, Jiang
;  APPLICANT: Charles, Boone
;  APPLICANT: Howard, Bussey
;  TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
;  FILE REFERENCE: 10182-005-999
;  CURRENT APPLICATION NUMBER: US/10/032,585
;  CURRENT FILING DATE: 2001-12-20
;  NUMBER OF SEQ ID NOS: 8000
;  SOFTWARE: Patentin version 3.1
SEQ ID NO 7646
;  LENGTH: 1881
;  TYPE: PRT
;  ORGANISM: Candida albicans
;  FEATURE:
;  NAME/KEY: MISC_FEATURE
;  LOCATION: (1881)..(1881)
;  OTHER INFORMATION: X-any amino acid
10-032-585-7646

Query Match     7.7%;   Score 212;   DB 7;   Length 1881;
Best Local Similarity   21.2%;   Pred. No. 4e-05;
Matches 135;   Conservative 115;   Mismatches 244;   Indels 144;   Gaps 23;

Qy    5 NSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGKK 64
          ::| |         | |       ::|     |::|| :  | :  | :    ::| :  |
Db  897 NTKKDYELQINNLNKS----NNEFKQKINELS---KKIESLTED-NKFNAKQLEEK 944

Qy   65 IHQNNGLDTEYNHN---GSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAARKKCSEFFTNK 121
          |||  ::  : | :     | | |  |  |      |:|| ||      |  :  |  :|
Db  945 LR-----DTEENNEHLMDKLRSASVAYNDLKKAKSES-----EEETVKAKEELETLTSK 993

Qy  122 ------------LKEKHTDLGK------EGVTDA------DAKEAILKTNGTKTKGAEELGKL 160
                      |:|:   |:|         | |  | :  || :| : ||: || : |:||
Db  994 IDNLEKELKEQQSKKNELEGQLQNITDSTNEKFKELEDELKSIKKSNKEISSQNSEL--- 1050

Qy  161 FESVEVLSKAAKEMLAN-------SVKELTSPVVAESPAMGSNSGKGGDSASTNPADE 211
          ||  |   |                |:| |  :  | |  | |  | |:|  |  |:|
Db 1051 --IQKLEKTEKDLQAKDEEIDKLKAETKSNIDNLNSEISSLQSKLKEAEESHSSTKDEH 1107

Qy  212 SAKGPNLTEISKKITDSNAFVL-------AVKEVETLVLSIDELAKKAIGQK--- 256
          :  ||| | ||||:||||| |         |:|||  :| :|:|:|||||  |
Db 1108 SSLSENLKKLKEEYENTKTSMIAKLSAKIEEHKKATDEIETKTHITDLQEEHAKQKSQF 1167

Qy  257 -IDNNGLAALNNQNGSLLAGAYAISTLITEKL---SKLKNLEELKTEIAKAKKCSEEFT 312
           |: |       ::| | | | |  |:   :|    |:|| ||| || :|:|| :||  |
Db 1168 ESERNDIKSNLDEANKELSDNREKLSNLEKEKTELNNKLKTQEEKISDLETSVAISEDKS 1227

Qy  313 NLKLKSGHADLGKQDATDDHAKAAILKTHATTDKGAKE------FKDLFESVEGLLKAA 364
          ::|   | |:||| |: :| | |  :|||       :|:      |:|| |||  ||||
Db 1228 KSLKHDIEDLKREKIKLE----TTLKENEETMFEKKEQLQVVNDCKELEACLKKLTETK 1283

Qy  365 QVALTNSVKELGHRNSNGGDSASTNPDESAKGPNLTV-ISKKITDSNAFLL------- 414
          :  | :   | |::  ||  :   :|  | :||| |   |||::||||| |
Db 1284 EKEINDLIRKL------EAAKSDHDTERKKLSLLIEDTKSESEKNVIKLNEQIEKLKG 1335

Qy  415 -AVKEVEALLSSI------DELSKAIGKKIIKNDGTLDNEANRNESLIAGAYEISKLITQ 466
           |:||| |:  :       |:|||::||:||:|| :|  :|  |:|:||   |  |::|:
Db 1336 EREKEVRDIQSQLAAKTTDWEKIKTLDKVLKEKS--DLEKTNKESVDTLKKEV----- 1387

Qy  467 KLSVLNSEELKKKKIKEAKDCSQKFTTKLKDSHAELGIQSVQDDNAKKAILKTHGTKDKGA 526
          :| :|||:||:|:|||||:|  :: | ||: |:|:|| :|   :|| |||||| | ::|
Db 1388 ------ENLKKEISLLEDQKKDDTTKYKELAAQLETKTSNLDSTTMELEKTELELKKVR 1440

Qy  527 KELEELFKSLESLS---------KAAQAALTNSVKEL 554
          :||:||||||:||          |||:||||||||:|
Db 1441 NELTEATSELTKLQDNNQSLTEEIEKTKAALTKSSKDL 1478
```

```
RESULT 14
US-10-072-851-15590
Sequence 15590, Application US/10072851
; GENERAL INFORMATION:
;  APPLICANT: Carr, Grant J.
;  APPLICANT: Xu, H. Howard
;  APPLICANT: Foulkes, J. Gordon
;  APPLICANT: Zamudio, Carlos
;  APPLICANT: Haselbeck, Robert
;  APPLICANT: Ohlsen, Karl L.
;  APPLICANT: Zyskind, Judith W.
;  APPLICANT: Wall, Daniel
;  APPLICANT: Trawick, John D.
;  APPLICANT: Yamamoto, Robert T.
;  APPLICANT: Roemer, Terry
;  APPLICANT: Jiang, Bo
;  APPLICANT: Boone, Charles
;  APPLICANT: Bussey, Howard
;  TITLE OF INVENTION: Methods for Identifying the Target of a Compound which Inhibit
;  TITLE OF INVENTION: Proliferation
;  FILE REFERENCE: ELITRA.028A
;  CURRENT APPLICATION NUMBER: US/10/072,851
;  CURRENT FILING DATE: 2002-02-08
;  PRIOR APPLICATION NUMBER: 60/267,636
;  PRIOR FILING DATE: 2001-02-09
;  NUMBER OF SEQ ID NOS: 15811
;  SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 15590
;  LENGTH: 1881
;  TYPE: PRT
;  ORGANISM: Candida albicans
;  FEATURE:
;  NAME/KEY: MISC_FEATURE
;  LOCATION: (1881)..(1881)
;  OTHER INFORMATION: X-any amino acid
US-10-072-851-15590

Query Match     7.7%;   Score 212;   DB 7;   Length 1881;
Best Local Similarity   21.2%;   Pred. No. 4e-05;
Matches 135;   Conservative 115;   Mismatches 244;   Indels 144;   Gaps 23;

Qy    5 NSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGKK 64
          ::| |         | |       ::|     |::|| :  | :  | :    ::| :  |
Db  897 NTKKDYELQINNLNKS----NNEFKQKINELS---KKIESLTED-NKFNAKQLEEK 944

Qy   65 IHQNNGLDTEYNHN---GSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAARKKCSEFFTNK 121
          |||  ::  : | :     | | |  |  |      |:|| ||      |  :  |  :|
Db  945 LR-----DTEENNEHLMDKLRSASVAYNDLKKAKSES-----EEETVKAKEELETLTSK 993

Qy  122 ------------LKEKHTDLGK------EGVTDA------DAKEAILKTNGTKTKGAEELGKL 160
                      |:|:   |:|         | |  | :  || :| : ||: || : |:||
Db  994 IDNLEKELKEQQSKKNELEGQLQNITDSTNEKFKELEDELKSIKKSNKEISSQNSEL--- 1050

Qy  161 FESVEVLSKAAKEMLAN-------SVKELTSPVVAESPAMGSNSGKGGDSASTNPADE 211
          ||  |   |                |:| |  :  | |  | |  | |:|  |  |:|
Db 1051 --IQKLEKTEKDLQAKDEEIDKLKAETKSNIDNLNSEISSLQSKLKEAEESHSSTKDEH 1107

Qy  212 SAKGPNLTEISKKITDSNAFVL-------AVKEVETLVLSIDELAKKAIGQK--- 256
          :  ||| | ||||:||||| |         |:|||  :| :|:|:|||||  |
Db 1108 SSLSENLKKLKEEYENTKTSMIAKLSAKIEEHKKATDEIETKTHITDLQEEHAKQKSQF 1167

Qy  257 -IDNNGLAALNNQNGSLLAGAYAISTLITEKL---SKLKNLEELKTEIAKAKKCSEEFT 312
           |: |       ::| | | | |  |:   :|    |:|| ||| || :|:|| :||  |
Db 1168 ESERNDIKSNLDEANKELSDNREKLSNLEKEKTELNNKLKTQEEKISDLETSVAISEDKS 1227

Qy  313 NLKLKSGHADLGKQDATDDHAKAAILKTHATTDKGAKE------FKDLFESVEGLLKAA 364
          ::|   | |:||| |: :| | |  :|||       :|:      |:|| |||  ||||
Db 1228 KSLKHDIEDLKREKIKLE----TTLKENEETMFEKKEQLQVVNDCKELEACLKKLTETK 1283

Qy  365 QVALTNSVKELGHRNSNGGDSASTNPDESAKGPNLTV-ISKKITDSNAFLL------- 414
          :  | :   | |::  ||  :   :|  | :||| |   |||::||||| |
Db 1284 EKEINDLIRKL------EAAKSDHDTERKKLSLLIEDTKSESEKNVIKLNEQIEKLKG 1335
```

```
RESULT 15
US-09-708-427-15046
Sequence 15046, Application US/09708427

GENERAL INFORMATION:
  APPLICANT: N. ALEXANDROV et al.
  TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
  TITLE OF INVENTION: THEREBY
  FILE REFERENCE: 2750-1243P
  CURRENT APPLICATION NUMBER: US/09/708,427
  CURRENT FILING DATE: 2000-11-09
  NUMBER OF SEQ ID NOS: 85364
  SOFTWARE: Patentin version 3.1
  SEQ ID NO 15046
  LENGTH: 1144
  TYPE: PRT
  ORGANISM: Arabidopsis thaliana
  FEATURE:
  NAME/KEY: misc_feature
  LOCATION: 1..1144
  OTHER INFORMATION: Xaa is any amino acid
  NAME/KEY: misc_feature
  LOCATION: 1..1144
  OTHER INFORMATION: Ceres Seq. ID 1828629
US-09-708-427-15046
```

```
Query Match      7.6%;
Best Local Similarity   21.8%;    Score 209;  DB 6;  Length 1144;
Matches  155;  Conservative 116;  Mismatches 247;  Indels 192;  Gaps 31;

Qy      2  ACNNSG---------------------------------------KDGNTSANSADESVKGPNLTEISKKITDSNAVL------  40
              ::|:|                                         :|| | :|| ||| || : ||||| :|| |
Db    102  SCTNSSVEEWKNKYHELEKEVEESNRSKSSASESME------------SVMKQLAELNHVLHETKSD                   156

Qy     41  ----------LAVKEVEALLSSIDEIA-----AKAIGKKI------------------------------HQNNG            70
                     |||| |||| |:||||| ||                                             ||
Db    157  NAAQKEKIELLEKTIEAQRTDLEEYGRQVCIAKEEASKLENLVESIKSELEISQEEKTRA                          216

Qy     71  LDTEYNHNGSLLAGAYAISTLLIKQKLD-GLKNEGLKEKIDAARKCSETFTNKLKEKHTDL                         129
           |:|                          :| ::|   || ||| ||   | |:|||
Db    217  LDNE------KAATSNIQNLLDQRTELSIELERCKVEEEKSKKDMESLTLALQEASTE-                           268

Qy    130  GKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKE----------ML---ANSV                        179
               || ::| |                  :|:|               :  |      :
Db    269  -----SSEAKATLLVCQ---------EELKNCESQVDSLKLASKETNEKYEKMLEDARNEI                         315

Qy    180  KELTSPVVA--ESPAMGSNSG------------KGGDSASTNPADESAKGPNLTEIS----                          222
              |  :  :|   : :||                 || | :| | |   |
Db    316  DSLKSTVDSIQNEFENSKAGWEQKELHLMGCVRKSEENSSQEEVSRLVNLLKESEEDA                            375

Qy    223  ----KKITDSNAFVLAVKEVETLVLSIDE-------LAKRAIGQKIDNNN---GLAALN                            267
                   :||     || | |::||            :| |||||      | | |
Db    376  CARKEEEASLKNNLKVAEGEVKYLQETLGEAKAESMKIKESLLDKEEDLKNVTAEISSLR                          435

Qy    268  NONGSLLAGAYAISTL---ITEKLSKLKNL----EELKTEIAKAKKCSEEFTNKLKSGHA                          320
               ::|  |  | | |      | | || |      || |:| |  |  |:|| |   :|
Db    436  EWEGSVLEKIEELSKVKESLVDKETKLQSITQEAEELKGREAAHMKQIEE----LSTANA                         491

Qy    321  DLGKQDATDDHAK--AAILKTHATTDKGA----KEFKDLFESVEGL---------------                          360
```
```

```
Qy    361  -LKAAQVALTNSVKELGHRNNSGGDSAS---TNPDESA-------------KGPNLTVISKK                         405
            :|          ||   | :: |          ||  :            ||||| | ||
Db    547  DLKEREVAYLKKIEELSVANESLVDKETKLQHIDQEAEELRGREAASHLKKIEELSKENEN                         606

Qy    406  ITDSNAFLLAVKEVE----------ALLSSIDELSKAIGKKIKNDGTLLDNEANRNE------                         451
            |:|                      ||:|  ||  |  |    ||  |  |
Db    607  LVDNVANMQNIAPESKDLREREVAYLKKIDELSTANGTLADNVTNLQNISEENKELRERE                          666

Qy    452  -SLIAGAYEISKLITQKLSVLNSEELKKKIKEAKDCSQKFTTKLIKDSHAELGIQSVQDDN                          510
             :|  |  || | || |  :|| |||| | | |   |   | | ||  ||||  ||  |
Db    667  TTLLKRAEELSEL--NESLVDKASKLQTVVQENEELRERETAYLKKIEELSKLHEILSDQ                          724

Qy    511  AKKAILKTHGTKDKGAKELEELF-KSLESLSKAAQALTNSVKELTNPVV                                     559
            |  :|     :|  |        ||:| || | | ||  ||| |
Db    725  ETKLQISNHEKEB--LKERETAYLKKIEELSK-VQEDLLNKENELHGMVV                                    771
```

```
Qy    415  -AVKEVEALLSSI--------DELSKAIGKKIKNDGTLDNEANRNESLLIAGAYEISKLITQ                         466
            ||||||                :|  |  || |  | ||:|| ::| | | | | | |||
Db   1336  EREKEVRDIQSQLAAKTIDWEKIKTTLDKVLKEKS--DLEKTNKESVDTLKKEV------                          1387

Qy    467  KLSVLNSEELKKKIKEAKDCSQKFTTKLKDSHAELGIQSVQDDNAKAILKTHGTKDKGA                           526
            || | ||  ||  |  |   |    | || ||                |:| | |   |
Db   1388  -------ENLKKEISLLEDQKDDTTKYKELAAQLETKTSNLDSTTMELEKTELELKKVR                          1440

Qy    527  KELEELFKSLESLS----------------KAAQAALTNSVKEL                                          554
            |  ||:  |  |:                 |    ||  || |
Db   1441  NELTEATSELTKLQDNNQSLTEEIEKTKAALFKSSKDL                                               1478
```

Search completed: March 18, 2002, 09:58:37
Job time: 395 sec

OM protein - protein search, using sw model

```
                    GenCore version 4.5
                    Copyright (c) 1993 - 2000 Compugen Ltd.

Run on:             March 18, 2002, 09:56:49 ; Search time 68.77 Seconds
                                              (without alignments)
                                              620.296 Million cell updates/sec

Title:              US-09-596-746a-52
Perfect score:      2750
Sequence:           1 MACNNSGKDGNTSANSADES.........KAAQAALTNSVKELTNPVVA 560

Scoring table:
                    BLOSUM62
                    Gapop 10.0 , Gapext 0.5
Searched:           219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters:     219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing:    Minimum Match 0%
                    Maximum Match 100%
                    Listing first 45 summaries

Database :          PIR_68:*
                    1: pir1:*
                    2: pir2:*
                    3: pir3:*
                    4: pir4:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

```
Result                     %
 NO.   Score   Query    Match  Length  DB    ID         Description
               Match
-------------------------------------------------------------------------------
  1     928     33.7            210     2   G70218    outer surface prot
  2     921     33.5            210     2   S69927    outer surface prot
  3     886     32.2            212     2   S20543    outer surface prot
  4     881     32.0            207     2   S37727    outer surface prot
  5     878     31.9            207     2   S69919    outer surface prot
  6     877     31.9            212     2   S69921    outer surface prot
  7     874     31.8            200     2   I40122    outer surface prot
  8     853     31.0            177     2   I40129    outer surface prot
  9     846     30.8            179     2   S54187    outer surface prot
 10     844     30.7            179     2   S54188    outer surface prot
 11     834.5   30.3            180     2   S54191    outer surface prot
 12     832     30.3            179     2   S54186    outer surface prot
 13     811     29.5            189     2   S70263    outer surface prot
 14     798     29.0            207     2   I40271    outer surface prot
 15     794     28.9            200     2   S54198    outer surface prot
 16     786     28.6            207     2   I40276    outer surface prot
 17     759.5   27.6            199     2   S54197    outer surface prot
 18     757.5   27.5            211     2   I40145    outer surface prot
 19     756.5   27.5            193     2   S70279    outer surface prot
 20     700     25.5            210     2   I40144    outer surface prot
 21     696.5   25.3            193     2   S70287    outer surface prot
 22     683.5   24.9            211     2   I40277    outer surface prot
 23     677.5   24.6            211     2   S69918    outer surface prot
 24     676.5   24.6            209     2   I40273    outer surface prot
 25     674.5   24.5            191     2   S70278    outer surface prot
 26     674.5   24.5            191     2   I40153    outer surface prot
 27     673.5   24.5            211     2   I40268    outer surface prot
 28     672.5   24.5            193     2   S70280    outer surface prot
 29     672     24.4            212     2   S69922    outer surface prot
 30     667     24.3            194     2   S70268    outer surface prot
 31     666.5   24.2            193     2   S70265    outer surface prot
 32     666.5   24.2            211     2   I40278    outer surface prot
 33     666     24.2            211     2   S70277    outer surface prot
 34     665.5   24.2            194     2   S69917    outer surface prot
 35     665     24.2            190     2   S70273    outer surface prot
 36     661     24.0            212     2   S70254    outer surface prot
 37     657.5   23.9            212     2   S70276    outer surface prot
 38     656.5   23.9            193     2   S70274    outer surface prot
 39     653.5   23.8            193     2   S54189    outer surface prot
 40     653.5   23.8            180     2   S70284    outer surface prot
 41     652.5   23.7            211     2   S69932    outer surface prot
 42     642.5   23.4            209     2   I40279    outer surface prot
 43     642     23.3            209     2   I40142    outer surface prot
 44     635     23.1            212     2   I40143    outer surface prot
 45     634.5   23.1            203     2   I40108    outer surface prot
```

ALIGNMENTS

```
RESULT: 1
G70218
outer surface protein C - Lyme disease spirochete
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 21-Jul-2000
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Accession: I40269
A;Molecule type: DNA
A;Residues: 1-210 <RES>
A;Cross-references: GB:D49497; NID:g707092; PIDN:BAA08457.1; PID:g769684
R;Frase r, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; Wh
ite,
son; D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vu
; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997
A;Accession: G70218
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Title: Outer surface protein C gene sequence analysis of Borrelia burgdorferi sensu
A;Reference number: I40269; MUID:96025162
A;Accession: I40269
A;Molecule type: DNA
A;Residues: 1-210 <JAU>
A;Cross-references: EMBL:X69596; NID:g311391; PIDN:CAA49306.1; PID:g311392
R;Livey, I.; Gibbs, C.P.; Schuster, R.; Dorner, F.
Mol. Microbiol. 18, 257-269, 1995
A;Title: Evidence for lateral transfer and recombination in OspC variation in Lyme di
A;Reference number: S70255; MUID:96296448
A;Accession: S70281
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 19-210 <LIV>
A;Cross-references: EMBL:L42887; NID:g858715; PIDN:AAB36995.1; PID:g1695212
A;Experimental source: strain Ip2
A;Gene: OspC
C;Superfamily: Lyme disease spirochete surface protein C
C;Genetics:
C;Gene: OspC
```

RESULT 2
S69927
outer surface protein C precursor - Lyme disease spirochete (strain PKa)
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
A;Variety: strain PKa
C;Date: 15-Feb-1997 #sequence_revision 27-Feb-1997 #text_change 26-May-2000
C;Accession: S69927; S72669
R;Jauris-Heipke, S.; Liegl, G.; Preac-Mursic, V.; Roessler, D.; Schwab, E.; Soutschek, E.
J. Clin. Microbiol. 33, 1860-1866, 1995
A;Title: Molecular analysis of genes encoding outer surface protein C (OspC) of Borre...
A;Reference number: I40047; MUID:95395018
A;Status: preliminary
A;Accession: S69927
A;Molecule type: DNA
A;Residues: 1-210 <JAU>
A;Cross-references: EMBL:X69589
A;Experimental source: strain PKa
A;Accession: S72669
A;Reference number: S72669
A;Molecule type: DNA
A;Residues: 1-124,'D',126-139,'E',141-210 <JAW>
C;...ne: ospc
C;...enetics:
C;...perfamily: Lyme disease spirochete surface protein C

```
Query Match          98.4%;  Score 928;  DB 2;  Length 210;
Best Local Similarity 33.7%;
Matches  188;  Conservative   2;  Mismatches  1;  Indels  0;  Gaps  0;

Qy    1  MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEVALLSSIDEIAKA   60
          :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   17  ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEVALLSSIDEIAKA   76

Qy   61  IGKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSEFFTN  120
          ||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db   77  IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSEFFTN  136

Qy  121  KLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK  180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  137  KLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK  196

Qy  181  ELTSPVVAESP  191
          |||||||||||
Db  197  ELTSPVVAESP  207
```

outer surface protein C precursor - Lyme disease spirochete
N;Alternate names: membrane-associated protein pC
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 21-Jul-2000
C;Accession: S20543; I40107; S69931; S70275; S34175
R;Fuchs, R.; Jauris, S.; Lottspeich, F.; Preac-Mursic, V.; Wilske, B.; Soutschek, E.
Mol. Microbiol. 6, 503-509, 1992
A;Title: Molecular analysis and expression of a Borrelia burgdorferi gene encoding a...
A;Reference number: S20543; MUID:92219995
A;Accession: S20543
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-212 <FUC>
A;Cross-references: EMBL:X62162; NID:g39992; PIDN:CAA44093.1; PID:g39393
R;Theisen, M.; Frederiksen, B.; Lebech, A.M.; Vuust, J.; Hansen, K.
J. Clin. Microbiol. 31, 2570-2576, 1993
A;Title: Polymorphism in ospC gene of Borrelia burgdorferi and immunoreactivity of Os...
A;Reference number: I40105; MUID:94075528
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-205 <RES>
A;Cross-references: EMBL:X73624; NID:g313273; PIDN:CAA52003.1; PID:g313274
R;Jauris-Heipke, S.; Liegl, G.; Preac-Mursic, V.; Roessler, D.; Schwab, E.; Soutschek
J. Clin. Microbiol. 33, 1860-1866, 1995
A;Title: Molecular analysis of genes encoding outer surface protein C (OspC) of Borre
A;Accession: S69931
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-212 <JAU>
A;Cross-references: EMBL:X69590
A;Experimental source: strain PWudi
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, February 1994
R;Livey, I.; Gibbs, C.P.; Schuster, R.; Dorner, F.
Mol. Microbiol. 18, 257-269, 1995
A;Title: Evidence for lateral transfer and recombination in OspC variation in Lyme di
A;Reference number: S70255; MUID:96296448
A;Accession: S70275
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 19-212 <LIV>
A;Cross-references: EMBL:L42883; NID:g858709; PIDN:AAB36991.1; PID:g1695209
A;Gene: ospc
C;Superfamily: Lyme disease spirochete surface protein C

RESULT 3
S20543

```
Query Match          97.4%;  Score 921;  DB 2;  Length 210;
Best Local Similarity 33.5%;
Matches  186;  Conservative   4;  Mismatches  1;  Indels  0;  Gaps  0;

Qy    1  MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEVALLSSIDEIAKA   60
          :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   17  ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEVALLSSIDEIAKA   76

Qy   61  IGKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSEFFTN  120
          ||||||||||||||:|||||||||||||||||:||||||||||||||||:||||||||||
Db   77  IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKOKLDGLKNEGLKEKIEAAKKCSEFFTN  136

Qy  121  KLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK  180
          |||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  137  KLKDKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK  196

Qy  181  ELTSPVVAESP  191
          |||||||||||
Db  197  ELTSPVVAESP  207
```

outer surface protein C precursor - Lyme disease spirochete (strain DK26)
A;Experimental source: strain DK26
R;Jauris-Heipke, S.; Liegl, G.; Preac-Mursic, V.; Roessler, D.; Schwab, E.; Soutschek
J. Clin. Microbiol. 33, 1860-1866, 1995
A;Title: Molecular analysis of genes encoding outer surface protein C (OspC) of Borre
A;Reference number: I40047; MUID:95395018
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-212 <JAU>
A;Cross-references: EMBL:X69590
A;Experimental source: strain JSB
A;Gene: ospc
C;Superfamily: Lyme disease spirochete surface protein C

RESULT 4
S37727

```
Query Match          90.7%;  Score 886;  DB 2;  Length 212;
Best Local Similarity 32.2%;
Matches  186;  Conservative   2;  Mismatches  3;  Indels  14;  Gaps  2;

Qy  195  SNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVEFLVLSIDELAKKAIG  254
          :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   20  NNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVEFLVLSIDELAKKAIG   79

Qy  255  QKIDNNNGLAALNNQNGSLLAGAYAISTLITEKLSKLKNLEELKTEIAKAKKCSEFFTNK  314
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   80  QKIDNNNGLAALNNQNGSLLAGAYAISTLITEKLSKLKNLEELKTEIAKAKKCSEFFTNK  139

Qy  315  LKSGHADLGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGLLKAAQVALTNSVKE  374
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  140  LKSGHADLGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGLLKAAQVALTNSVKE  199

Qy  375  LGHRNNSGGDSASTNP--DESAKGP  397
          |:|    |||   || |
Db  200  L-----------TSPVVAESPKKP  212
```

outer surface protein C precursor - Lyme disease spirochete
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 26-May-2000
C;Accession: S37727
R;Jauris-Helpke, S.; Fuchs, R.; Motz, M.; Preac-Mursic, V.; Schwab, E.; W
Med. Microbiol. Immunol. 182, 37-50, 1993
A;Title: Genetic heterogeneity of the genes coding for the outer surface protein C (OspC)
A;Reference number: S37726; MUID:93268136
A;Accession: S37727
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-207 <JAU>
A;Cross-references: EMBL:X69595; NID:g311393; PIDN:CAA49305.1; PID:g311394
C;Superfamily: Lyme disease spirochete surface protein C

```
Qy   559 VA 560
         ||
Db   200 VA 201

Qy   499 AELGIQSVQDDNAKKAILKTHGTKDKGAKELEELFKSLESLSKAAQAALTNSVKELTNPV 558
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   140 AELGIQSVQDDNAKKAILKTHGTKDKGAKELEELFKSLESLSKAAQAALTNSVKELTNPV 199

Qy   439 NDGTLDNEANRNESLIAGAYEISKLITQKLSVLNSEELKKKIKEAKDCSOKFTTKLKDSH 498
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    80 NDGTLDNEANRNESLIAGAYEISKLITQKLSVLNSEELKKKIKEAKDCSOKFTTKLKDSH 139

Qy   379 NNSGGDSASTNPDESAKGPNLTVISKKITDSNAFLLAVKEVEALLSSIDELSKAIGKKIK 438
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    20 NNSGGDSASTNPDESAKGPNLTVISKKITDSNAFLLAVKEVEALLSSIDELSKAIGKKIK 79
```

Query Match   32.0%; Score 881; DB 2; Length 207;
Best Local Similarity 99.5%; Pred. No. 1.8e-32;
Matches 181; Conservative   1; Mismatches   0; Indels   0; Gaps   0;

RESULT 5
S69919
outer surface protein C precursor - Borrelia garinii
A;Species: Borrelia garinii
A;Variety: strain PTrob
C;Date: 06-Dec-1996 #sequence_revision 14-Feb-1997 #text_change 26-May-2000
C;Accession: S69919
R;Jauris-Helpke, S.; Liegl, G.; Preac-Mursic, V.; Roessler, D.; Schwab, E.; Soutschek, F
J. Clin. Microbiol. 33, 1860-1866, 1995
A;Title: Molecular analysis of genes encoding outer surface protein C (OspC) of Borrella
A;Reference number: I40047; MUID:95395018
A;Accession: S69919
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-207 <JAU>
A;Experimental source: strain PTrob
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1995
A;Cross-references: EMBL:X83554; NID:g872027; PIDN:CAA58544.1; PID:g872028
F;19-207/Product: outer surface protein C #status predicted <MAT>
F;1-18/Domain: signal sequence #status predicted <SIG>
C;Superfamily: Lyme disease spirochete surface protein C
A;Gene: ospc

Query Match   31.9%; Score 878; DB 2; Length 207;
Best Local Similarity 98.9%; Pred. No. 2.4e-32;
Matches 180; Conservative   2; Mismatches   0; Indels   0; Gaps   0;

```
Qy   379 NNSGGDSASTNPDESAKGPNLTVISKKITDSNAFLLAVKEVEALLSSIDELSKAIGKKIK 438
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    20 NNSGGDSASTNPDESAKGPNLTVISKKITDSNAFLLAVKEVEALLSSIDELSKAIGKKIK 79

Qy   439 NDGTLDNEANRNESLIAGAYEISKLITQKLSVLNSEELKEKIKEAKDCSEKFTTKLKDSH 498
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    80 NDGTLDNEANRNESLIAGAYEISKLITQKLSVLNSEELKEKIKEAKDCSEKFTTKLKDSH 139
```

Qy   499 AELGIQSVQDDNAKKAILKTHGTKDKGAKELEELFKSLESLSKAAQAALTNSVKELTNPV 558
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   140 AELGIQSVQDDNAKKAILKTHGTKDKGAKELEELFKSLESLSKAAQAALTNSVKELTNPV 199

Qy   559 VA 560
         ||
Db   200 VA 201

RESULT 6
S69921
outer surface protein C precursor - Borrelia afzelii (strain PBO)
A;Species: Borrelia afzelii
A;Variety: strain PBO
C;Date: 06-Dec-1996 #sequence_revision 14-Feb-1997 #text_change 26-May-2000
C;Accession: S69921; I40047
R;Jauris-Helpke, S.; Liegl, G.; Preac-Mursic, V.; Roessler, D.; Schwab, E.; Soutschek, F
J. Clin. Microbiol. 33, 1860-1866, 1995
A;Title: Molecular analysis of genes encoding outer surface protein C (OspC) of Borre
A;Reference number: I40047; MUID:95395018
A;Accession: S69921
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-212 <JAU>
A;Cross-references: EMBL:X81521; NID:g804954; PIDN:CAA57241.1; PID:g804955
A;Experimental source: strain PBO
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, September 199
A;Genetics:
A;Gene: ospC
C;Superfamily: Lyme disease spirochete surface protein C

Query Match   31.9%; Score 877; DB 2; Length 212;
Best Local Similarity 89.8%; Pred. No. 2.7e-32;
Matches 184; Conservative   4; Mismatches   3; Indels 14; Gaps   2;

```
Qy   195 SNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETLVLSIDELAKKAIG 254
         :||||||||||||:|||||||||||||:||||||||||||||||||||||||||||||||
Db    20 NNSGKGGDSASTDPADESAKGPNLTEISKKITDSNAFVLAVKEVETLVLSIDELAKKAIG 79

Qy   255 QKIDNNNGLAALNNQNGSLLAGAYAISTLITEKLSKLKNLEELKTEIAKAKKCSEEFTNK 314
         ||||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||
Db    80 QKIDNNNGLAALNNQNGSLLAGAYAISTLITEKLSKLKNEELKTEIAKAKKCSEEFTNK 139

Qy   315 LKSGHADLGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGLLKAAQVALTNSVKE 374
         ||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||
Db   140 LKSGHADLGKQDATDEHAKAAILKTHATTDKGAKEFKDLFESVEGLLKAAQVALTNSVKE 199

Qy   375 LGHRNNSGGDSASTNP--DESAKGP 397
         ||:|:|       ||:|||
Db   200 L--------------TSPVVAESPKKP 212
```

RESULT 7
I40122
outer surface protein C - Lyme disease spirochete (strain DK6) (fragment)
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 26-May-2000
A;Accession: I40122
R;Theisen, M.; Frederiksen, B.; Lebech, A.M.; Vuust, J.; Hansen, K.
J. Clin. Microbiol. 31, 2570-2576, 1993
A;Title: Polymorphism in ospC gene of Borrelia burgdorferi and immunoreactivity of Os
A;Reference number: I40105; MUID:94075528
A;Accession: I40122; S34177
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-200 <RES>
A;Cross-references: EMBL:X73626; NID:g313277; PIDN:CAA52005.1; PID:g313278
A;Genetics:
A;Gene: ospC
C;Superfamily: Lyme disease spirochete surface protein C

RESULT   8
I40129
outer surface protein C - Lyme disease spirochete (strain TXGW) (fragment)
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 26-May-2000
C;Accession: I40129; S54199
R;Thelsen, M.; Borre, M.; Mathiesen, M.J.; Lebech, A.M.; Hansen, K.
J. Bacteriol. 177, 3036-3044, 1995
A;Title: Evolution of the Borrelia burgdorferi outer surface protein OspC.
A;Reference number: I40104; MUID:95286481
A;Accession: I40129
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-177 <RES>
A;Cross-references: EMBL:X84783; NID:g793825; PIDN:CAA59254.1; PID:g793826
C;Genetics:
A;Gene: ospC
C;Superfamily: Lyme disease spirochete surface protein C

```
Query Match            31.8%;  Score 874; DB 2; Length 200;
Best Local Similarity  98.9%;  Pred. No. 3.5e-32;
Matches 179;  Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy  379 NNSGGDSASTNPDESAKGPNLTVISKKITDSNAFLLAVKEVEALLSSIDELSKAIGKKIK 438
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   20 NNSGGDSASTNPDESAKGPNLTVISKKITDSNAFLLAVKEVEALLSSIDELSKAIGKKIK 79

Qy  439 NDGTLDNEANRNESLIAGAYEISKLLITQKLSVLNSEELKKIKEAKDCSQKFTTKLKDSH 498
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   80 NDGTLDNEANRNESLIAGAYEISKLLITQKLSVLNSEELKKIKEAKDCSEKFTTKLKDSH 139

Qy  499 AEELGQSVQQDDNAKKAILKTHGTKDKGAKELEELFKSLESLSKAAQQAALTNSVKELTNPV 558
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  140 AEELGQSVQQDDNAKKAILKTHGTKDKGAKELEELFKSLESLSKAAQQAALTNSVKELTNPV 199

Qy  559 ---                                          559
Db  200 ---                                          200
```

RESULT   9
S54187
outer surface protein C - Borrelia afzelii (strain DK2) (fragment)
C;Species: Borrelia afzelii
C;Date: 08-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 08-Dec-2000
C;Accession: I40115; S54187
R;Thelsen, M.; Borre, M.; Mathiesen, M.J.; Mikkelsen, B.; Lebech, A.M.; Hansen, K.
J. Bacteriol. 177, 3036-3044, 1995
A;Title: Evolution of the Borrelia burgdorferi outer surface protein OspC.
A;Reference number: I40104; MUID:95286481
A;Accession: I40115
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-179 <THE>
A;Cross-references: EMBL:X84766; NID:g793801; PIDN:CAA59237.1; PID:g793802
A;Experimental source: strain DK2
A;Note: submitted to the EMBL Data Library, February 1995
C;Genetics:
A;Gene: ospC
C;Superfamily: Lyme disease spirochete surface protein C

```
Query Match            31.0%;  Score 853; DB 2; Length 177;
Best Local Similarity  98.9%;  Pred. No. 2.5e-31;
Matches 175;  Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy   11 NTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSSIDEIAAKAIGKKIHONG 70
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 NTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGKKIHQNG 60

Qy   71 LDTEYNHNGSLLAGAYAISTLIKQKLDGILKNEGLKEKIDAAKKCSEFTTNKLKEKHTDLG 130
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 LDTENNHNGSLLAGAYAISTLIKQKLDGILKNEGLKEKIDAAKKCSEFTTNKLKEKHTDLG 120

Qy  131 KEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKELTSPVV 187
            |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db  121 KRGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKELTSPVV 177
```

RESULT   10
S54188
outer surface protein C - Borrelia afzelii (strain DK3) (fragment)
C;Species: Borrelia afzelii
C;Date: 08-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 08-Dec-2000
C;Accession: I40116; S54188
R;Thelsen, M.; Borre, M.; Mathiesen, M.J.; Mikkelsen, B.; Lebech, A.M.; Hansen, K.
J. Bacteriol. 177, 3036-3044, 1995
A;Title: Evolution of the Borrelia burgdorferi outer surface protein OspC.
A;Reference number: I40104; MUID:95286481
A;Accession: I40116
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-179 <THE>
A;Cross-references: EMBL:X84771; NID:g793803; PIDN:CAA59242.1; PID:g793804
A;Experimental source: strain DK3
A;Note: submitted to the EMBL Data Library, February 1995
C;Genetics:
A;Gene: ospC
C;Superfamily: Lyme disease spirochete surface protein C

```
Query Match            30.8%;  Score 846; DB 2; Length 179;
Best Local Similarity  99.4%;  Pred. No. 5.1e-31;
Matches 173;  Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy  202 DSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETLVLSIDELAKKAIGQKIDNNN 261
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 DSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETLVLSIDELAKKAIGQKIDNNN 60

Qy  262 GLAALNNQNGSLLAGAYAISTLITTEKLSKLKNLEELKTEIAKAKKCSEEFTNKLKSGHAD 321
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 GLAALNNQNGSLLAGAYAISTLITEKLSKLKNLEELKTEIAKAKKCSEEFTNKLKSGHAD 120

Qy  322 LGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGLLKAAQVALTNSVKEL 375
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 LGKQDATDDHAKAVILKTHATTDKGAKEFKDLFESVEGLLKAAQVALTNSVKEL 174
```

RESULT   11
S54191
outer surface protein C - Borrelia afzelii (strain DK8) (fragment)
C;Species: Borrelia afzelii
C;Date: 08-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 08-Dec-2000
C;Accession: I40119; S54191
R;Thelsen, M.; Borre, M.; Mathiesen, M.J.; Mikkelsen, B.; Lebech, A.M.; Hansen, K.
J. Bacteriol. 177, 3036-3044, 1995
A;Title: Evolution of the Borrelia burgdorferi outer surface protein OspC.

```
Query Match            30.7%;  Score 844; DB 2; Length 179;
Best Local Similarity  99.4%;  Pred. No. 6.3e-31;
Matches 173;  Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy  202 DSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETLVLSIDELAKKAIGQKIDNNN 261
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 DSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETLVLSIDELAKKAIGQKIDNNN 60

Qy  262 GLAALNNQNGSLLAGAYAISTLITTEKLSKLKNLEELKTEIARAKKCSEEFTNKLKSGHAD 321
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 GLAALNNQNGSLLAGAYAISTLITEKLSKLKNLEELKTEIARAKKCSEEFTNKLKSGHAD 120

Qy  322 LGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGLLKAAQVALTNSVKEL 375
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 LAKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGLLKAAQVALTNSVKEL 174
```

A;Reference number: I40119; MUID:95286481
A;Accession: I40119
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-180 <THE>
A;Cross-references: EMBL:X84776; NID:g793809; PIDN:CAA59247.1; PID:g793810
A;Note: submitted to the EMBL Data Library, February 1995
C;Molecule type: DNA
C;Genetics:
A;Gene: ospc
C;Superfamily: Lyme disease spirochete surface protein C

Query Match   30.3%;   Score 834.5;   DB 2;   Length 180;
Best Local Similarity   98.9%;   Pred. No. 1.7e-30;
Matches 173;   Conservative   1;   Mismatches   1;   Indels   1;   Gaps   1;

Qy   202   DSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETLVLSIDELAKKAIGQKIDNNN 261
         |||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||
Db   1   DSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETLVLSIDELAKKAIGQKIDNNN 60

Qy   262   GLAALNNQNGSLLAGAYAIS-TLITEKLSKLKNLEELKTEIAKAKKCSEEFTNKLKSGHA 320
         |||||:|||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db   61   GLAALHNQNGSLLAGAYAISTTLITEKLSKLKNLEELKTEIAKAKKCSEEFTNKLKSGHA 120

Qy   321   DLGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGLLKAAQVALTNSVKEL 375
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121   DLGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGLLKAAQVALTNSVKEL 175

RESULT   12
S54186
outer surface protein C - Borrelia afzelii (strain DK21) (fragment)
C;Species: Borrelia afzelii
C;Date: 08-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 08-Dec-2000
C;Accession: I40114; S54186
R;Thelsen, M.; Borre, M.; Mathiesen, M.J.; Mikkelsen, B.; Lebech, A.M.; Hansen, K.
J. Bacteriol. 177, 3036-3044, 1995
A;Reference number: I40114
A;Title: Evolution of the Borrelia burgdorferi outer surface protein OspC.
A;Accession: I40114
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-179 <THE>
A;Cross-references: EMBL:X84767; NID:g793799; PIDN:CAA59238.1; PID:g793800
A;Note: submitted to the EMBL Data Library, February 1995
C;Genetics:
A;Gene: ospC
C;Superfamily: Lyme disease spirochete surface protein C

Query Match   30.3%;   Score 832;   DB 2;   Length 179;
Best Local Similarity   97.7%;   Pred. No. 2.1e-30;
Matches 170;   Conservative   4;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   202   DSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETLVLSIDELAKKAIGQKIDNNN 261
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1   DSASTSPADESAKGPNLTKISKKITDSNAFVLAVKEVETLVLSIDELAKKAIGQKIDNDN 60

Qy   262   GLAALNNQNGSLLAGAYAVAISTLITEKLSKLKNLEELKTEIAKAKKCSEEFTNKLKSGHAD 321
         |||||||||||||||||:|||||||||||:|||:|||||:||||||||||||||||||||||
Db   1   GLAALNNQNGSLLAGAYAVAISTLITEKLSKLKNLEELKTEIAEAKKCSEEFTNKLKSGHAD 60

Qy   322   LGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGLLKAAQVALTNSVKEL 375
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121   LGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGLLKAAQVALTNSVKEL 174

RESULT   13
S70263
outer surface protein C - Lyme disease spirochete (strain Kl11) (fragment)
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
A;Variety: strain Kl11
C;Date: 12-Feb-1998 #sequence_revision 26-Feb-1998 #text_change 26-May-2000
C;Accession: S70263; S77616
R;Livey, I.; Gibbs, C.P.; Schuster, R.; Dorner, F.
Mol. Microbiol. 18, 257-269, 1995
A;Title: Evidence for lateral transfer and recombination in OspC variation in Lyme di
A;Accession: S70263
A;Status: nucleic acid sequence not shown
A;Cross-references: EMBL:L42880; NID:g858707; PIDN:AAB36989.1; PID:g1695207
C;Genetics:
A;Gene: ospC
C;Superfamily: Lyme disease spirochete surface protein C

Query Match   29.5%;   Score 811;   DB 2;   Length 189;
Best Local Similarity   90.7%;   Pred. No. 1.9e-29;
Matches 165;   Conservative   11;   Mismatches   6;   Indels   0;   Gaps   0;

Qy   379   NNSGGDSASTNPDESAKGPNLTVISKKITDSNAFLLAVKEVEALLSSIDELSKAIGKKIK 438
         |||||:|||||||||||||||||||||||||||:|:|||||||||||||||||||||||:|
Db   2   NNSGGDTASTNPDESAKGPDLTVISKKITDSNAVVLVVKEVEALLSSIDELSKAIGKKIR 61

Qy   439   NDGTLDNEANRNESLIAGAYEISKLITQKLSVLNSEELKKKIKEAKDCSQKFTTKLKDSH 498
         ||||||||||||||||||||||||||||||||||||:|||||||||||||:|||||||||
Db   62   NDGTLDNEANRNESLIAGAYEISKLITQKLSVLNSEELKEKIKEAKDCSQKFTTKLKDSH 121

Qy   499   AELGIQSVQDDNAKKAILKTHGTDKDKGAKELEELFKSLESLSKAAQAALTNSVKELTNPV 558
         |||||:|||:|:|||:|||||||||:|||||||:||:||||||||||||| |||||||:||
Db   122   AELGIQNVQDDNAKRAILKTHGNKDKGAKELSESLEKLSKAAQAALANSVQELTSPV 181

Qy   559   VA 560
         ||
Db   182   VA 183

RESULT   14
I40271
outer surface protein C precursor - Borrelia garinii
C;Species: Borrelia garinii
J. Clin. Microbiol. 33, 2415-2420, 1995
R;Fukunaga, M.; Hamase, A.
A;Title: Outer surface protein C gene sequence analysis of Borrelia burgdorferi sensu
A;Reference number: I40269; MUID:96025162
C;Date: 04-Sep-1997 #sequence_revision 04-Sep-1997 #text_change 26-May-2000
C;Accession: I40271
A;Accession: I40271
A;Status: Preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-207 <RES>
A;Cross-references: GB:D49377; NID:g1041103; PIDN:BAA08375.1; PID:g1041104
C;Superfamily: Lyme disease spirochete surface protein C

Query Match   29.0%;   Score 798;   DB 2;   Length 207;
Best Local Similarity   88.5%;   Pred. No. 7.9e-29;
Matches 161;   Conservative   15;   Mismatches   6;   Indels   0;   Gaps   0;

Qy   379   NNSGGDSASTNPDESAKGPNLTVISKKITDSNAFLLAVKEVEALLSSIDELSKAIGKKIK 438
         |||||:|:|||||||||||:|:|||||||||||:||:|||||||||||||||||||||:||
Db   20   NNSGGDTASTNPDESAKGPNLIEISKKITDSNAVVLAVKEVEALLSSIDELSKAIGKKIR 79

Qy   439   NDGTLDNEANRNESLIAGAYEISKLITQKLSVLNSEELKKKIKEAKDCSQKFTTKLKDSH 498
         ||||||||||||||||||||||||||:||:||||||:|:|||||:||:|||||||:|||||
Db   80   NDGTLDNEANRNESLIAGAYEISKLITQKLSVLNSEELKEKIKEAKDCSEKFTTKLRDSH 139

```
Qy    499  AELGIQSVQDDNAKKAILKTHGTKDKGAKELEELFKSLESLSKAAQQAALTNSVKELTNPV  558
           ||||:||:|||||||||||||||||||:|| ||||||||:|||||||  :||||||||||
Db    140  AELGVQNVQDDNAKRAILKTHGNKDKGAKELKELSESLENLAKAAQOAASSNSVKELTSPV  199

Qy    559  VA  560
           ||
Db    200  VA  201
```

RESULT 15
S54198
outer surface protein C - Borrelia garinii (strain SL20) (fragment)
C;Species: Borrelia garinii
C;Date: 08-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 08-Dec-2000
C;Accession: I40126; S54198
R;Theisen, M.; Borre, M.; Mathiesen, M.J.; Lebech, A.M.; Hansen, K.
bacteriol. 177, 3036-3044, 1995
A;Title: Evolution of the Borrelia burgdorferi outer surface protein OspC.
A;Reference number: I40104; MUID:95286481
A;Accession: I40126
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-200 <THE>
A;Cross-references: EMBL:X84781; NID:g793819; PIDN:CAA59252.1; PID:g793820
A;Note: submitted to the EMBL Data Library, February 1995
C;Genetics:
A;Gene: ospC
C;Superfamily: Lyme disease spirochete surface protein C

Query Match   28.9%; Score 794; DB 2; Length 200;
Best Local Similarity   89.0%; Pred. No. 1.1e-28;
Matches 161; Conservative 13; Mismatches   7; Indels   0; Gaps   0;

```
Qy    379  NNSGGDSASTNPDESAKGPNLTVISKKITDSNAFLLAVKEVEALLSSIDELSKAIGKKIK  438
           ||||| |||||||:|:|||||||||||||||| :| |||||||||:||||||||||||
Db     20  NNSGGDTASTNPDESAKGPDLTVISKKITDSNAVVLVVKEVEALLSSIDELSKAIGKKIR   79

Qy    439  NDGTLDNEANRNESLIAGAYEISKLITQKLSVLNSEELKKKIKEAKDCSQKFTTKLKDSH  498
           ||||||||||||||||||||||||||:|||||||||||||:|||||||||:|||| ||
Db     80  NDGTLDNEANRNESLIAGAYEISKLITQKLSVLNSEELKEKIKEAKDCSEKFTTKLRXSH  139

Qy    499  AELGIQSVQDDNAKKAILKTHGTKDKGAKELEELFKSLESLSKAAQQAALTNSVKELTNPV  558
           ||||:||:|||||||||||||||||||:|| ||||||||:|||||||  :||||||||||
Db    140  AELGVQNVQDDNAKRAILKTHGNKDKGAKELKELSESLEKLSKAAQOAALANSVQELTSPV  199
```

```
OM protein - protein search, using sw model

                          Copyright (c) 1993 - 2000 Compugen Ltd.
                          GenCore version 4.5

Run on:         March 18, 2002, 10:11:57 ; Search time 39.62 Seconds
                                            (without alignments)
                                            518.231 Million cell updates/sec

Title:          US-09-596-746A-52
Perfect score:  2750
Sequence:       1 MACNNSGKDGNTSANSADES.........KAAQAALTNSVKELTNPVVA 560

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters:    100059

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :      SwissProt_39:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                              SUMMARIES

Result           %
No.     Score   Match  Length  DB  ID
------  -----   -----  ------  --  ----------
   1      928   33.7     210   1   OSC1_BORBU
   2      886   32.2     212   1   OSC2_BORBU
   3    402.5   14.6     214   1   VM24_BORHE
   4      398   14.5     215   1   VM03_BORHE
   5    210.5    7.7     1957  1   YD86_SCHPO
   6    207.5    7.5     1790  1   USO1_YEAST
   7    207.5    7.5     2869  1   RBP1_PLAVB
   8    204.5    7.4     1509  1   MYSN_ACACA
   9    193.5    7.0     2017  1   MYSN_DROME
  10    193.5    7.0     2116  1   MYS2_DICDI
  11    193      7.0     1433  1   REST_CHICK
  12    192      7.0     1966  1   MYSB_CAEEL
  13    189      6.9     1022  1   SCA4_RICPR
  14    189      6.9     1427  1   REST_HUMAN
  15    187      6.8      875  1   ZIP1_YEAST
  16    187      6.8     1164  1   BAG_STRAG
  17    186      6.8     1167  1   CAGA_HELPJ
  18    185      6.7     1189  1   SCI1_CHICK
  19    185      6.7     1251  1   RBP2_PLAVB
  20    184.5    6.7     1636  1   BUD3_YEAST
  21    183.5    6.7     1189  1   YJH6_YEAST
  22    183.5    6.7     1940  1   MYH3_CHICK
  23    182.5    6.6     1786  1   LMB1_MOUSE
  24    180.5    6.6      775  1   YHG6_BACSU
  25    180      6.5     3672  1   LM12_CAEEL
  26    179.5    6.5     2411  1   MYSA_DROME
  27    179      6.5     1251  1   MYH3_RAT
  28    178.5    6.5      727  1   MFP1_ARATH
  29    178.5    6.5     1290  1   XCPC_XENLA
  30    177.5    6.4     1875  1   LMB1_HUMAN
  31    177      6.4     1786  1   MLP1_YEAST
  32    177      6.4     2748  1   NUM1_YEAST
  33    176.5    6.4      998  1   SCA4_RICAK
  34    176.5    6.4     1935  1   MYH7_RAT
  35    176.5    6.4     1976  1   MYHA_BOVIN
  36    176      6.4     1947  1   MYSC_CAEEL
  37    175.5    6.4      692  1   MYS_PODCA
  38    175.5    6.4     1935  1   MYH7_HUMAN
  39    175      6.4     1038  1   CIN8_YEAST
  40    174.5    6.3      722  1   MFP1_TOBAC
  41    174.5    6.3      600  1   ABRA_PLAFG
  42    174      6.3     1939  1   MYH6_MESAU
  43    173.5    6.3     1391  1   MST2_DROHY
  44    173.5    6.3     1935  1   MYH7_PIG
  45    173.5    6.3     1972  1   MYHB_HUMAN
```

```
DB  ID          Description
--  ----------  -----------
1   Q07337      borrelia bu
1   Q08137      borrelia bu
1   P32778      borrelia he
1   Q02448      borrelia he
1   Q10411      schizosacch
1   P25386      saccharomyc
1   Q00798      plasmodium
1   P05659      acanthamoeb
1   Q99323      drosophila
1   P08799      dictyostell
1   O42184      gallus gall
1   P02566      caenorhabdi
1   Q92d49      rickettsia
1   P30622      homo sapien
1   P31111      saccharomyc
1   P27951      streptococc
1   O2z1t1      helicobacte
1   Q90988      gallus gall
1   Q00799      plasmodium
1   P25558      saccharomyc
1   P47035      saccharomyc
1   P02565      gallus gall
1   P02469      mus musculu
1   P32399      bacillus su
1   Q21313      caenorhabdi
1   P05661      drosophila
1   P12847      rattus norv
1   O1w85       arabidopsis
1   P50532      xenopus lae
1   P07942      homo sapien
1   P02565      saccharomyc
1   Q00402      saccharomyc
1   Q9aix9      rickettsia
```

RESULT
    1
OSC1_BORBU

ID   OSC1_BORBU   STANDARD;   PRT;   210 AA.
AC   Q07337;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   OUTER SURFACE PROTEIN C PRECURSOR (PC).
GN   OSPC OR BBB19.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OG   Plasmid lp54.
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 35210 / B31;
RX   MEDLINE=93239332; PubMed=8478108;
RA   Wilske B., Preace-Mursic V., Jauris S., Pradel I., Soutschek E.,
RA   Schwab E., Wanner G.;
RT   "Immunological and molecular polymorphisms of OspC, an immunodominant
RT   major outer surface protein of Borrelia burgdorferi.";
RL   Infect. Immun. 61:2182-2191(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 35210 / B31;
RX   MEDLINE=93268136; PubMed=8098841;
RA   Jauris-Heipke S., Fuchs R., Motz M., Preac-Mursic V., Schwab E.,
RA   Will G., Wilske B.;
RT   "Genetic heterogeneity of the genes coding for the outer surface
RT   protein C (OspC) and the flagellin of Borrelia burgdorferi.";
RL   Med. Microbiol. Immunol. 182:37-50(1993).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 35210 / B31;
RX   MEDLINE=94041630; PubMed=8225587;
RA   Padula S.J., Sampieri A., Dias F., Szczepanski A., Ryan R.W.;
RT   "Molecular characterization and expression of p23 (OspC) from a North
RT   American strain of Borrelia burgdorferi.";
RL   Infect. Immun. 61:5097-5105(1993).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 35210 / B31;
RX   MEDLINE=96025162; PubMed=7494039;
RA   Fukunaga M., Hamase A.;
RT   "Outer surface protein C gene sequence analysis of Borrelia
RT   burgdorferi sensu lato isolates from Japan.";
RL   J. Clin. Microbiol. 33:2415-2420(1995).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 35210 / B31;
RX   MEDLINE=98065943; PubMed=9403685;
RA   Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA   Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
RA   Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
RA   Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,
RA   van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,

ALIGNMENTS
```

RA   Utterback T., Watthey L., McDonald L., Artiach P., Bowman C.,
RA   Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
RA   Smith H.O., Venter J.C.;
RT   "Genomic sequence of a Lyme disease spirochaete, Borrelia
RT   burgdorferi.";
RL   Nature 390:580-586(1997).
CC   -!- FUNCTION: NOT KNOWN; MAJOR IMMUNODOMINANT PROTEIN.
CC   -!- SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID
CC   ANCHOR.
CC   ---
DR   EMBL; X69596; CAA49306.1; -.
DR   EMBL; U01894; AAA16058.1; -.
DR   EMBL; D49947; BAA08457.1; -.
DR   EMBL; AE000792; AAC66329.1; -.
DR   TIGR; BBB19; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
DR   PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KW   Outer membrane; Lipoprotein; Signal; Plasmid; Antigen;
KW   Complete proteome.
FT   SIGNAL   1   18   BY SIMILARITY.
FT   CHAIN   19   210   OUTER SURFACE PROTEIN C.
FT   LIPID   19   19   N-ACYL DIGLYCERIDE (BY SIMILARITY).
SQ   SEQUENCE   210 AA; 22340 MW; 7A4FC978F91777BF CRC64;

Query Match   33.7%;
Best Local Similarity   98.4%;
Matches 188; Conservative   2; Mismatches   1; Indels   0; Gaps   0;

Qy   1 MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA 60
            :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   17 ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA 76

Qy   61 IGKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKCSEFTFN 120
            |||||||||||||| ||||||||||||||||||||||||||||||| |||||||||||
Db   77 IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSEFTFN 136

Qy   121 KLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 180
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   137 KLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 196

Qy   181 ELTSPVVAESP 191
            |||||||||||
Db   197 ELTSPVVAESP 207

RESULT 2
OSC2_BORBU
ID   OSC2_BORBU   STANDARD;   PRT;   212 AA.
AC   Q08137;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   OUTER SURFACE PROTEIN C PRECURSOR (PC).
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
GN   OSPC.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX   MEDLINE=92219995; PubMed=1560779;
RA   Fuchs R., Jauris S., Lottspeich F., Preac-Mursic V., Wilske B.,

---

RA   Soutschek E.;
RT   "Molecular analysis and expression of a Borrelia burgdorferi gene
RT   encoding a 22 kDa protein (pC) in Escherichia coli.";
RL   Mol. Microbiol. 6:503-509(1992).
RN   [2]
RP   SEQUENCE OF 1-205 FROM N.A.
RC   STRAIN=DK26;
RX   MEDLINE=94075528; PubMed=8253951;
RA   Theisen M., Frederiksen B., Lebech A.M., Vuust J., Hansen K.;
RT   "Polymorphism in ospC gene of Borrelia burgdorferi and
RT   immunoreactivity of OspC protein: implications for taxonomy and for
RT   use of OspC protein as a diagnostic antigen.";
RL   J. Clin. Microbiol. 31:2570-2576(1993).
CC   -!- FUNCTION: NOT KNOWN; MAJOR IMMUNODOMINANT PROTEIN.
CC   -!- SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID
CC   ANCHOR.
CC   ---
DR   EMBL; X62162; CAA44093.1; -.
DR   EMBL; X73624; CAA52003.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
KW   Outer membrane; Lipoprotein; Signal; Plasmid; Antigen;
FT   SIGNAL   1   18   BY SIMILARITY.
FT   CHAIN   19   212   OUTER SURFACE PROTEIN C.
FT   LIPID   19   19   N-ACYL DIGLYCERIDE (BY SIMILARITY).
SQ   SEQUENCE   212 AA; 22499 MW; C206C231FBF2E7D4 CRC64;

Query Match   32.2%;
Best Local Similarity   90.7%;
Matches 186; Conservative   2; Mismatches   3; Indels   14; Gaps   2;

Qy   195 SNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETLVLSIDELAKKAIG 254
            :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   20 NNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETLVLSIDELAKKAIG 79

Qy   255 QKIDNNGLAALNQNGSLLAGAYAISTLITTEKLSKLKNLEELKTEIAKAKKCSEEFTNK 314
            |||||||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db   80 QKIDNNGLAALNNQNGSLLAGAYAISTLITTEKLSKLKNLEELKTEIAKAKKCSEEFTNK 139

Qy   315 LKSGHADLGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGLLKAAQVALTNSVKE 374
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   140 LKSGHADLGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGLLKAAQVALTNSVKE 199

Qy   375 LGHRNSGGDSASTNP--DESAKGP 397
            | | :| |||
Db   200 L------------------TSPVVAESPKKP 212

RESULT 3
VM24_BORHE
ID   VM24_BORHE   STANDARD;   PRT;   214 AA.
AC   P32778;
DT   01-OCT-1993 (Rel. 27, Created)
DT   01-OCT-1993 (Rel. 27, Last sequence update)
DT   01-FEB-1994 (Rel. 28, Last annotation update)
DE   VARIABLE MAJOR OUTER MEMBRANE LIPOPROTEIN 24 PRECURSOR.
GN   VMP24.
OS   Borrelia hermsii.
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OG   Plasmid.
OX   NCBI_TaxID=140;
RN   [1]
RP   SEQUENCE FROM N.A.

Column 1

```
RC   STRAIN=SSP, HS1 SEROTYPE 24;
RX   MEDLINE=93133110; PubMed=1484486;
RA   Restrepo B.I., Kitten T., Carter C.J., Infante D., Barbour A.G.;
RT   "Subtelomeric expression regions of Borrelia hermsii linear plasmids
RT   are highly polymorphic.";
RL   Mol. Microbiol. 6:3299-3311(1992).
CC   -!- FUNCTION: SERVES TO AVOID THE HOST IMMUNE RESPONSE BY CHANGING
CC       FROM ONE SURFACE EXPOSED VMP TO ANOTHER.
CC   -!- SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID
CC       ANCHOR.
CC   -!- SIMILARITY: STRONG, TO VMP3.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; L04786; AAA22964.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
DR   PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KW   Outer membrane; Lipoprotein; Signal; Plasmid.
FT   SIGNAL        1     18
FT   CHAIN        19    214       VARIABLE MAJOR OUTER MEMBRANE
FT                                LIPOPROTEIN 24.
FT   LIPID        19     19       N-ACYL DIGLYCERIDE (PROBABLE).
SQ   SEQUENCE   214 AA;  22541 MW;  F1583F510246F7C7 CRC64;
```

```
Query Match    14.6%;   Score 402.5;  DB 1;   Length 214;
Best Local Similarity  44.9%;  Pred. No. 4.9e-11;
Matches  92;  Conservative  37;  Mismatches  59;  Indels  17;  Gaps  5;

Qy    1 MACNNSGKDGNTSANSADESVKGP----NLTEISKKITDSNAVLLAVKEVEALLSSIDEI 56
          ||||| ||| :||                 ||||||||||||||||||| ||||| ::|:
Db   17 MSCNNNGGPE----LKSDEVAKSDGTVLDLAKVSKKIKEASAFAASVKEVETLVKSVDEL 71

Qy   57 AAKAIGKKIHQONGLDTEYNHNGSLLAGAYAISTLIKOKL------DGLKNEGLKEKIDA 110
          |||| || :|||||| |  |||:|||||||||||:|||           |||| || ||
Db   72 -AKAIGKKIKNDGGLDTEAGQNGSLIAGVHSVSVAVKIKVGALETTSGISNE-LKTKITE 129

Qy  111 AKKCSETFTNKLKEKHTDLGKEGGVTDADAKEAILKNGTKTKGAEELGKLFESVEVLSKA 170
          |||||||:||||||| ||| |||| |: || ||||| |||| |  || |||||||||||
Db  130 VKSKAEAFLNKLKDGHTFELGKKDASDDDTKRAIKKDNSDKTKGASELEALNTAVDALLKA 189

Qy  171 AKEMLANSVKELTSPVVAESPAMGS 195
          |||||||||||||| |||||||||
Db  190 AEGEVEAAIKELTAPVKAEKPSQNN 214
```

Column 2

```
RT   are highly polymorphic.";
RL   Mol. Microbiol. 6:3299-3311(1992).
CC   -!- FUNCTION: SERVES TO AVOID THE HOST IMMUNE RESPONSE BY CHANGING
CC       FROM ONE SURFACE EXPOSED VMP TO ANOTHER.
CC   -!- SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID
CC       ANCHOR.
CC   -!- SIMILARITY: STRONG, TO VMP24.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; L04789; AAA22967.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
DR   PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KW   Outer membrane; Lipoprotein; Signal; Plasmid.
FT   SIGNAL        1     18
FT   CHAIN        19    215       VARIABLE MAJOR OUTER MEMBRANE
FT                                LIPOPROTEIN 3.
FT   LIPID        19     19       N-ACYL DIGLYCERIDE (PROBABLE).
SQ   SEQUENCE   215 AA;  23139 MW;  684C74D35F87C771 CRC64;
```

```
Query Match    14.5%;   Score 398;  DB 1;   Length 215;
Best Local Similarity  45.4%;  Pred. No. 7.5e-11;
Matches  93;  Conservative  35;  Mismatches  61;  Indels  16;  Gaps  5;

Qy    1 MACNNSGKDGNTSANSADESVKGP----NLTEISKKITDSNAVLLAVKEVEALLSSIDEI 56
          :|||| |||  |||               ||||||||||||||||||| ||||| ::|:
Db   17 MSCNNNGGPE----LKSDEVAKSDGTVLDLAKISKKIKIDASDFAASVKEVHTLVKSIDEL 71

Qy   57 AAKAIGKKI-HQONGLDTEYNHNGSLLAGAYAISTLIKOKLDGLK------NEGLKEKIDA 110
          ||||| ||   |||||| |  |||:|||||| | ||||||| |||        |||| ||
Db   72 -AKAIGKKIKNDNSNFEDENDHNGSLIAGVFQVILTVKAKLTSLEQIGISDELKTEVGM 130

Qy  111 AKKCSETFTNKLKEKHTDLGKEGGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKA 170
          |||||||:||||||| ||| |||| |: || ||||| |||| |  || |||||||||||
Db  131 VKKESEAFVTQVKSKHTDLAKEGVTDAHAKSAILVTDGTKDKGAAELIKLNTAIDELLKA 190

Qy  171 AKEMLANSVKELTSPVVAESPAMGS 195
          |||||||||||||| ||||||||||
Db  191 ANDAVETVIKELTASVKAEKPSQNN 215
```

```
RESULT 4
ID   VM03_BORHE    STANDARD;    PRT;    215 AA.
AC   O02448;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   01-FEB-1994 (Rel. 28, Last annotation update)
DE   VARIABLE MAJOR OUTER MEMBRANE LIPOPROTEIN 3 PRECURSOR.
GN   VMP3.
OS   Borrelia hermsii.
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=140;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SSP, HS1 SEROTYPE 3;
RX   MEDLINE=93133110; PubMed=1484486;
RA   Restrepo B.I., Kitten T., Carter C.J., Infante D., Barbour A.G.;
RT   "Subtelomeric expression regions of Borrelia hermsii linear plasmids
```

Column 3

```
RESULT 5
ID   YD86_SCHPO    STANDARD;    PRT;    1957 AA.
AC   Q10411;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   HYPOTHETICAL 222.8 KDA PROTEIN C1F3.06C IN CHROMOSOME I.
GN   SPACF3.06C.
OS   Schizosaccharomyces pombe (Fission yeast).
OC   Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC   Schizosaccharomycetales; Schizosaccharomycetaceae;
OC   Schizosaccharomyces.
OX   NCBI_TaxID=4896;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=972;
RA   Connor R., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;
RL   Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
```

```
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb.sib.ch).
CC
DR   EMBL; Z70690; CAA94624.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE 1957 AA; 222785 MW; 3F480CA0671D59DA CRC64;
```

Query Match   7.7%; Score 210.5; DB 1; Length 1957;
Best Local Similarity 22.4%; Pred. No. 0.065;
Matches 157; Conservative 116; Mismatches 218; Indels 209; Gaps 35;

```
Qy      8 KDGNTSANSADE---SVKGPNLTEISKKITDSNAVLLAV-KEVEALLSSIDEIAAKAIGK   63
          |:||           :||  ||||:|| |||||||  |||  ||| ||||: |||:|
Db    360 KDSRTSNSQLEEMVELKESNRT-IHSQLTDAESKLSSFEQENKSLKGSIDE-------     410

Qy     64 KIHQNN------------GLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLK--EKIDA-A  111
          |||||                :|||| | | | ||||||  |||| |||   || | |||| |
Db    411 --YQNNLSSKDKMVKQVSSQLEEARSSLAHATGKLAEINSERD-FQNKKIKDFEKIEQDL    467

Qy    112 KKCSETFTNKLKEKHDILGKEGVTDADAKEAILKTNGTKTFGAEELGKLFESV--EVLSK   169
          :|  | | :|||||| :| |  | ||  ||:|  ||||| | | |||| |:|::: :|||
Db    468 RACLNSSSNELKEKSALIDKKDOELNNLREQIKE---QKKVSESTQSSLQSLQRDILNE     523

Qy    170 AAK-EMLANSVKELTSPVVAESPAMGSNSGKGGDSASTNPADESAKGPNLTEISKK----   224
          :|  | ||:| |:|   |  |  |  |:| | | |  |:||||:|  | |  | :|
Db    524 KKKHEVYESQLNELKGELQTEI-----SNSEHLSSQLSTLAAEKEAAVATNNELSESKNSL   579

Qy    225 ITDSNAF---------VLAVKEVETLVLSIDELAKK--AIGQKIDNNNGL-----       263
          | :|| |         | | |:    ||:|  :|    |||:||:|| ||
Db    580 QTFLCNAFQEKLAKSVWQLKENEQNTSSLDTSFKLNESHQELENNHQTITKQLDTSSKL    639

Qy    264 -----------AALNNQNGSLLAGAYAISTLITEKLSKL---------               291
                      ||| | :  || | |: | |  |  | ::|
Db    640 QQLQLERANFEQKESTLSDENNDL-------RTKLIKLEESNKSLIKKQEDVDSLE       688

Qy    292 KNLEELKTEIAKAK------------KCSEEFTNKLKSGHADIGKQ-----DATDDHAKAA  335
          ||  | | |  |                :|||| |:|  |||:|     |  : | |:|
Db    689 KNIQTLKEDLRKSEEALRFSKLEAKNLREVIDNLKGKHETLEAQRNDLHSSLSDAKNTNA   748

Qy    336 ILKTHATTDKGAKEEKDLFESVEGL--------LKAAQVALTNSVKELGH-----       377
          || :  | || :  ||| |:|  | |        | |||: |:| | ::|
Db    749 ILSSELT--KSSEDVKRLFANVELLTQDSKAMKQSFTSLVNSYQSISNLYHELRDDHVNM   806

Qy    378 --RNNS--GGDSASTNPDESAKGPNLTVISKKITDSNAFLLAVKEV--EALLSSIDELSK   431
            | ||   | | ||   |:| | | |  |::||||| | || ||:  ||| ||||:| |
Db    807 QSQNNTLLESESKLKTDCENLTQQNMTLI-----DNVQKLMHKHVNQESKVSELKEVNG    860

Qy    432 AIGKKIKN-DGTLDNEANRNESLLIAGAYEISK------LITQKLSVLNS-------     473
          || :|| | || || | |:|| | | |  |||        ||:||  ::|
Db    861 KLSLDLKNLRSSLNVAISDNDQIITQLAELSKNYDSLEQESAQLNSGILKSLEAEKQLLHT   920

Qy    474 --EEL--------KKKIKEAK--DCSQKFTTKLKDSHAELGIQSVQDDN--AKKAILKT    518
            |||        :||| |     |  :| | |  || ||  |: :   | |:| |||
Db    921 ENEELHIRLDKLITGKLKIEESKSSDLGKKLTARQEE-----ISNLKENMSQSQAITSV    974

Qy    519 HGTKDKGAKELEELFKSLESLSRAAQAALTNSVKELTNPV                     558
          : |:||| | | ||::| |: |:  | |: |::|:|
Db    975 KSKLD-----ETLSKSSK--LEADIEHLKNKV                             999
```

```
RESULT 6
ID   USO1_YEAST     STANDARD;      PRT;   1790 AA.
AC   P25386;
DT   01-MAY-1992 (Rel. 22, Created)
DT   01-MAY-1992 (Rel. 22, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   INTRACELLULAR PROTEIN TRANSPORT PROTEIN USO1.
GN   USO1 OR INT1 OR YDL058W.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91185402; PubMed=2010462;
RA   Nakajima H., Hirata A., Ogawa Y., Yonehara T., Yoda K.,
RA   Yamasaki M.;
RT   "A cytoskeleton-related gene, uso1, is required for intracellular
RT   protein transport in Saccharomyces cerevisiae.";
RL   J. Cell Biol. 113:245-260(1991).
RN   [2]
RP   SEQUENCE OF 782-1790 FROM N.A.
RA   Hostetter M.K., Herman D.J., Bendel C.M., McClellan M., Tao N.,
RA   Kendrick K.E.;
RL   Submitted (FEB-1993) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 1-8 FROM N.A.
RA   Bai Y., Synington L.S.;
RL   Submitted (MAY-1996) to the EMBL/GenBank/DDBJ databases.
-!- FUNCTION: REQUIRED FOR PROTEIN TRANSPORT FROM THE ER TO THE GOLGI
    COMPLEX.
-!- SUBCELLULAR LOCATION: CYTOPLASMIC. ASSOCIATED WITH INTRACELLULAR
    MEMBRANES. PROBABLY PRESENT ON VESICLES OPERATIONAL BETWEEN THE
    ER AND THE GOLGI COMPLEX.
-!- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, COMPOSED
    OF AN HEPTAPEPTIDE REPEAT PATTERN CHARACTERISTIC OF ALPHA-HELICAL
    COILED COILS. MAY FORM FILAMENTOUS STRUCTURES IN THE CELL.
-!- SIMILARITY: BELONGS TO THE VDP/USO1/YBL047C FAMILY.
--------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
DR   EMBL; X54378; CAA38253.1; -.
DR   EMBL; L03188; AAB00143.1; -.
DR   EMBL; U53668; AAB66659.1; -.
DR   PIR; A38455; A38455.
DR   HSSP; P80220; 1DIP.
DR   SGD; S0002216; USO1.
DR   InterPro; IPR002017; Spectrin.
KW   Transport; Protein transport; Golgi stack; Cytoskeleton; Coiled coil.
FT   DOMAIN        1    724       GLOBULAR HEAD.
FT   DOMAIN      725   1790       COILED COIL (POTENTIAL).
FT   DOMAIN        1    465       CHARGED (HYPER-HYDROPHILIC).
FT   DOMAIN      487    991       DISPENSABLE FOR THE PROTEIN FUNCTION.
FT   DOMAIN     1172   1786       ASP/GLU-RICH (ACIDIC).
FT   CONFLICT    847    847       G -> E (IN REF. 2).
FT   CONFLICT    924    924       E -> K (IN REF. 2).
FT   CONFLICT   1253   1253       V -> I (IN REF. 2).
FT   CONFLICT   1319   1319       I -> V (IN REF. 2).
FT   CONFLICT   1461   1461       N -> S (IN REF. 2).
FT   CONFLICT   1581   1581       N -> S (IN REF. 2).
FT   CONFLICT   1600   1600       I -> V (IN REF. 2).
FT   CONFLICT   1661   1661       R -> S (IN REF. 2).
FT   CONFLICT   1772   1772       D -> DEEDDEE (IN REF. 2).
SQ   SEQUENCE 1790 AA; 206424 MW; 6CE2B21E69FD4818 CRC64;
```

Query Match   7.5%; Score 207.5; DB 1; Length 1790;
Best Local Similarity 21.4%; Pred. No. 0.079;
Matches 143; Conservative 104; Mismatches 259; Indels 161; Gaps 22;

```
Qy      3 CNNSGKDG------------NTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSS   52
          ||| | |             ||:|||||||||||||||||||||||||||||||| ||||||
Db    929 CNNLSKEKEHISKELVEYKSRPQSHQNLV--AKLTEKLKSLANNYKDMQA-ENESLIKA     984

Qy     53 IDEIAAKAIGKKIHQNNGLD--TEYNHNGSLLAGAY--------AISTL------IKQK    95
```

```
Db   985  VEESKNESSIQLSNLQNKIDSMSQERENFQIERGSIEKNIEQLKKTISDLEQTKEIISK 1044
          ::|:            :|:      :|:
Qy    96  LDGLKNEG------LKEKIDAAKKCSETFTNKLKE----------KHTDLGKEGVTD 136
          |||:|        |:||          :|           :|||
Db  1045  SDSSKDEYESQTSLLKEKLETATTANDENVKRISELTKTREELEALEAAYKNLKNELETK 1104

Qy   137  ADAKEAILK----------TNGTKTK-----GAEELGKLFESVEVLSKAA 171
          |:|:|             |:|       :|||:||     :
Db  1105  LETSEKALKEVKENEEHLKEEKIQLEKEATETKQQLNSLRANLESLEKEHEDLAAQLKKY 1164

Qy   172  KEMLAN-------------SVKELTSPVVAESPAMGSNSNGKGGDSASTN 207
          ||:||               ||:|
Db  1165  EEQIANKEROVNEEISQLNDEITSTOQENESTKKKNDELEGEVKAMKSTS---EEQSNLK 1221

Qy   208  PADESAKGPNLTEISKKITDSNAFVL-AVKEVETLVLSIDEL------AKKAIGOKIDNN 260
          ::         :           :||   ||            :||
Db  1222  KSEIDALNLQIKELKLKKNETNEASLLESIKSVESETVKIKELQDECNFKEKEVSELEDKL 1281

Qy   261  NGLAALNNQNGSLLAGAYAISTLITEKLSKLKNLEELKTEIAKAKKCSEEFTNKLKSGHA 320
          ||:|              ||::||         :||        :|:|
Db  1282  KASEDKNSKYLELQKESEKIKEELDAKTELKIQLEKITNLSKAKSESELSRLK----- 1337

Qy   321  DLGKQDATDDHAKAALLKTHATTDKGAKEFKDLFESVEGLLKAAQVALTNSVKELGHRNN 380
          |                :|:
Db  1338  -KT-----SSEERKNAEEQLEKLKNEIQIRNQAFEKERKLLN- 1373

Qy   381  SGGDSASTNPDESAKGPNLTVISKKITDSNAFLLAVKEVEALLSSIDELSKAIGKKIKND 440
          :|:|:|||                :        |:      :|| :
Db  1374  ---EGSSTITQEYSEKIN-TLEDELIRLQNENELKAKEIDNTRSELEKVS-----LSND 1423

Qy   441  GTLDNEANRNESLIAGAYEISKLITOKLSVLNS---------EELKKKIKEAKDCSQKFT 491
          ||:                             ::           :|
Db  1424  ELLEEKQNTIKSLQDEILSYKDKITRNDEKLLSIERDNKRDLESLKEQLRAAQESRAKVE 1483

Qy   492  TKLKDSHAELGTQSVQDDNAKKAILKTHGTKDRGAKEL-----EELFKSLESLSKAAQAAL 547
          |:|                   :|     ::::           |||
Db  1484  EGLKKLEEESSKEKAELEKSKEMMKKLESTIESNETELKSSMETIRKSDEKILEQSKKSA- 1542

Qy   548  TNSVKEL 554
          ::| |
Db  1543  EEDIKNL 1549
```

```
RESULT 7
RBP1_PLAVB

ID   RBP1_PLAVB     STANDARD;     PRT;   2869 AA.
AC   Q00798;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-APR-1993 (Rel. 25, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   RETICULOCYTE BINDING PROTEIN 1 PRECURSOR.
GN   RBP1.
OS   Plasmodium vivax (strain Belem).
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=31273;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92315338; PubMed=1617731;
RA   Galinski M.R., Medina C.C., Ingravallo P., Barnwell J.W.;
RT   "A reticulocyte-binding protein complex of Plasmodium vivax
RT   merozoites.";
RL   Cell 69:1213-1226(1992).
CC   -!- FUNCTION: INVOLVED IN RETICULOCYTE ADHESION. SPECIFICALLY BINDS TO
CC       HUMAN RETICULOCYTE CELLS.
CC   -!- SUBUNIT: HOMODIMER (POTENTIAL).
CC   -!- SUBCELLULAR LOCATION: MEMBRANE-BOUND.
```

```
DR   EMBL; M88097; AAA29743.1; -.
DR   HSSP; P36956; 1AM9.
KW   Malaria; Receptor; Signal; Transmembrane.
FT   SIGNAL        1     17       POTENTIAL.
FT   CHAIN        18   2869       RETICULOCYTE BINDING PROTEIN 1.
FT   DOMAIN       18   2807       EXTRACELLULAR.
FT   TRANSMEM   2808   2826       POTENTIAL.
FT   DOMAIN     2827   2869       CYTOPLASMIC.
FT   SITE       1030   1032       CELL ATTACHMENT SITE (POTENTIAL).
FT   SITE       2599   2601       CELL ATTACHMENT SITE (POTENTIAL).
SQ   SEQUENCE   2869 AA;  330213 MW;  B9DBE44220SEBCFF CRC64;

Query Match          7.5%;  Score 207.5;  DB 1;  Length 2869;
Best Local Similarity  20.4%;  Pred. No. 0.13;
Matches  127;  Conservative 116;  Mismatches 260;  Indels 121;  Gaps 24;
```

```
Qy    25  NLTEISKKIT-DSNAVLLAVKEVEALLSSIDE------IAAKAIGKKIHQNNG----LDT 73
          |:           ||  ||:|||||||      ::
Db  1141  HINTVAHGITSNKNEILKSVKEVEDKLNIVEQNEDYKVKVNPENEKQLEAIRGSMSKLKE 1200

Qy    74  EYNHNGSLLAGAYAISTLIKOKLDGLKNEGLKERKIIDAAKKCSETFTNKLLKEKHTDLGKEG 133
          :|            :|:          :|:       :|:
Db  1201  VINKHVSEMTQLESTANTILKSNARGKENEHDLEELNKTGQOMRDIYELKIKKIAPEEL-KEG 1259

Qy   134  VTDADAKEAILKTNGTKTRGAEE-LGKLFESVEV----LSKAAKEMLANSVKELTSPVVA 188
          |||:|:|       |:           :         |||
Db  1260  TVN-ELKDANEKANKVEPEPERNIIGHVLERITVEKDRAGKVVEEM--NSLKIKIEKLIQ 1316

Qy   189  ESPAMGSNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETLVLSIDEL 248
          |:                         :|
Db  1317  ET-----------SDDSQNELVTTSITPKHLENAKGYEDVIKRNEEDSIQLREK 1358

Qy   249  AKKAIGQKIDNNNGLAALNNQNGSLLAGAYAISTLITEKLSKLKNLEEL-----KTEIA 302
          |:|             |:|||:                   :|:         ||:|
Db  1359  AKSL--ETLDEMKRLVQQVNMN--LQSAIQGNAGISKELNELKGVIELLISTNYSSILE 1413

Qy   303  KAKKCSEEFTNKLKSGHADLGKQDATDDHAKAAI-----LKTHATTDKGAKEFKDLFESV 357
          ||                   |:|||||||||       ||
Db  1414  YVKKNSSESVRFSQLANGEFTKAEGEEKNASARLAEEARLKEQIVKLDYSDIDDKVKKI 1473
```

```
RESULT 8
MYSN_ACACA

ID   MYSN_ACACA     STANDARD;     PRT;   1509 AA.
AC   P05659;
DT   01-NOV-1988 (Rel. 09, Created)
DT   01-NOV-1988 (Rel. 09, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   MYOSIN II HEAVY CHAIN, NON MUSCLE.
```

```
Qy   358  EG-----LKRAAQVALT-----NSVKEL--GHRNN---------------SGGDSASTNPDESA 394
          ||       ||:                |::                  :|
Db  1474  EGIKREILKMKESALTFWEESERKFQMCSSHMENAKEGKKKIEYLKNNGDGGKANITDSQ 1533

Qy   395  KGPNLTVISK------------KITDSNAFLLAVKEV-----EALLSSI------D 427
          ||                    :|||||| ||||:        |||||
Db  1534  MEEVGNYVSKAEHAFHTVEAQVDRKTKAFCESIVAYVTKMDNLFNESLMKEVKVKCEKKND 1593

Qy   428  ELSKAIGKKIKNDGTLDNEANRNESLIAGAYEISKLITOKLSVLNSEELKKKIKEAKDCS 487
          |:                 |:|       ||::|
Db  1594  EAEKYSAKLKPYDGRIKARVSENERKISELKEKAKVEKKESSQLN-DVSTKSLLQIDNCR 1652

Qy   488  OKFTTKLKD-----------SHAELGIOSV---QDDNAKKAILKTHGTRDKGAKELEEL 532
          |:               |:|        :|||
Db  1653  QQLDSVLSNIGRVKQNALQYFDSADKSMKSVLPISELGAEKSLDKVKAAKESYERNLETV 1712

Qy   533  FKSLESLSKAAQAALTNSVKELTN 556
          :|:                ||:
Db  1713  QNEMSRIN-VEEGSLTIDIDKKITD 1735
```

OS   Acanthamoeba castellanii (Amoeba).
OC   Eukaryota; Acanthamoebidae; Acanthamoeba.
OX   NCBI_TaxID=5755;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87308395; PubMed=3040773;
RA   Hammer J.A. III, Bowers B., Paterson B.M., Korn E.D.;
RT   "Complete nucleotide sequence and deduced polypeptide sequence of a
RT   nonmuscle myosin heavy chain gene from Acanthamoeba: evidence of a
RT   hinge in the rodlike tail.";
RL   J. Cell Biol. 105:913-925(1987).
CC   -!- FUNCTION: MYOSIN IS A PROTEIN THAT BINDS TO F-ACTIN & HAS ATPASE
CC   ACTIVITY THAT IS ACTIVATED BY F-ACTIN.
CC   -!- SUBUNIT: MYOSIN II HEAVY CHAIN IS TWO-HEADED. IT SELF-ASSEMBLES
CC   INTO FILAMENTS. HEXAMER OF 2 HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI
CC   LIGHT CHAIN SUBUNITS (MLC) AND 2 REGULATORY LIGHT CHAIN SUBUNITS
CC   (MLC-2).
CC   -!- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CC   CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CC   CHARACTERISTIC OF ALPHA-HELICAL STRUCTURES. THIS REGION IS
CC   INTERRUPTED BY A HINGE AND JOINED BY A NONHELICAL TAILPIECE WHERE
CC   THE REGULATORY PHOSPHORYLATION SITES RESIDE.
CC   -!- MISCELLANEOUS: THE HINGE REGION MAY PLAY A KEY ROLE IN MEDIATING
CC   THE EFFECT OF HEAVY CHAIN PHOSPHORYLATION ON ENZYMATIC ACTIVITY.
CC   -!- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC   -!- SIMILARITY: CONTAINS 2 IQ DOMAINS.
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --
DR   EMBL; Y00624; CAA68663.1; -.
DR   PIR; A27224; A27224.
DR   HSSP; P08799; 1MND.
DR   InterPro; IPR000048; IQ.
DR   InterPro; IPR001609; myosin_head.
DR   Pfam; PF00612; IQ; 2.
DR   Pfam; PF00063; myosin_head; 1.
DR   PRINTS; PR00193; MYOSINHEAVY.
DR   ProDom; PD000355; myosin_head; 1.
DR   SMART; SM00015; IQ; 1.
DR   SMART; SM00242; MYSc; 1.
DR   PROSITE; PS50096; IQ; 1.
KW   Myosin; Coiled coil; Actin-binding; ATP-binding; Calmodulin-binding;
KW   Methylation; Alkylation; Phosphorylation; Multigene family.
FT   DOMAIN   1   789   MYOSIN HEAD-LIKE.
FT   DOMAIN   790   819   IQ.
FT   DOMAIN   848   1509   COILED COIL (POTENTIAL).
FT   DOMAIN   848   1226   ALPHA-HELICAL TAILPIECE (S2).
FT   DOMAIN   1227   1252   HINGE.
FT   DOMAIN   1253   1509   LIGHT MEROMYOSIN (LMM).
FT   DOMAIN   1253   1482   ALPHA-HELICAL TAILPIECE (LMM).
FT   DOMAIN   1483   1509   NONHELICAL TAILPIECE (LMM).
FT   NP_BIND   182   189   ATP.
FT   DOMAIN   660   682   ACTIN-BINDING.
FT   DOMAIN   766   780   ACTIN-BINDING.
FT   MOD_RES   133   133   METHYLATION (TRI-) (POTENTIAL).
FT   MOD_RES   700   700   ALKYLATION (SH-1) (POTENTIAL).
FT   MOD_RES   1489   1489   PHOSPHORYLATION.
FT   MOD_RES   1494   1494   PHOSPHORYLATION.
FT   MOD_RES   1499   1499   PHOSPHORYLATION.
SQ   SEQUENCE   1509 AA; 171201 MW; 2CE49BE51173D17E CRC64;

RESULT 9
MYSN_DROME
ID   MYSN_DROME   STANDARD;   PRT;   2017 AA.
AC   Q99323;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   MYOSIN HEAVY CHAIN, NON-MUSCLE (ZIPPER PROTEIN) (MYOSIN II).
GN   ZIP.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RX   MEDLINE=90349606; PubMed=2117279;
RA   Ketchum A.S., Stewart C.T., Stewart M., Kiehart D.P.;
RT   "Complete sequence of the Drosophila nonmuscle myosin heavy-chain
RT   transcript: conserved sequences in the myosin tail and differential
RT   splicing in the 5' untranslated sequence.";
RL   Proc. Natl. Acad. Sci. U.S.A. 87:6316-6320(1990).
CC   -!- FUNCTION: NONMUSCLE MYOSIN APPEARS TO BE RESPONSIBLE FOR
CC   CELLULARIZATION. REQUIRED FOR MORPHOGENESIS AND CYTOKINESIS.
CC   -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
CC   SHORT FORM, ARE PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 IQ DOMAIN.
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its

Qy   28 EISKKITDSNAVLLAVKEVEA------LLSSIDE-------IAAKAIGKKIHQNGLDTEYNH   77
              : | | | | :||||| | | | :   :|| | : |: |:
Db   1379 ---DRDSIERRNRDAE-----------RQVRDLRAQLDDALSRLDSEKRA----   1414

Qy   78 NGSLLAGAYAISTLIKQKLDGLIKNEGLKEKIDA-- -AKKCSETFINKLKEKHTDLGKEGV 134
              : | | || | || :| | || || || ||| | | :| | | |
Db   1415 -KEKSVEANRELKKVVLDRERQSLESLSK   1442

Qy   135 TDA------DAKEAILKTNGTKTKGAEELGKLFESVELSKAAKEMLANSVKELTSPVVAESPA 192
Db   994 EA------AAHDSLKKKEEDLSRE-LRETKDALADAENISETLRSKL--KNTERGADDV   1043

Qy   193 MGSNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETILVLSIDELAKKA 252
Db   934 ELQEETSASNDILEQKKLEAEKGELKASLEEEERNRKALQEEAKTKVESERNELQDKYED 993

Qy   253 IGQKI-DNNGLAALNNQNGSLLAGAYAISTLIITEKLSKLKNLEELKTEIAKAKKCSEEF   311
Db   1044 RNELDDVTATKLQLEKTKKSLEEELAQTRAQLEE-EKSGKEAASSKAKQLGQOLEDARSE 1102

Qy   312 TNKLKS---GHADLGKQDAIDDHAKAAIL-----KTHATTDKGAKEFKDLFESVEGL   360
Db   1103 VDSLKSK--LSAAEKSLKTAKDQN-RDLDEQLEDE--------RTVRANVDK-QKKA   1147

Qy   361 LKAAQVALTNSVKELGHRN------NSGGDSASTNPDESAKGPNLTIVISKKI-TDSNAFL 413
Db   1148 LEAKLTELEDQVTALDGQOKNAAAAQAKTLKTQVDETKRRLEEAEASAAARLEKERKNALDE 1207

Qy   414 LAVKEVE----------ALLSSIDELSKAIGKKIKN-DGTLDNEAN---RNESLIAGAY   458
Db   1208 VAQLTADLIDAERDSAQQORRKLNTRISELQSELENAPKTGGASSEEVKRLEGELRLEEE 1267

Qy   459 EISKLITQKLSVLNSEELKKIKEAKDCSQKFTTKLKDSHAELGIQSVQDDNAKKAILKT 518
Db   1268 LITAQEARAAAEKNLDKANLELEELRQEADDAARDNDKLVKD------NRKLKADLDEAR 1321

Qy   519 HGTKDKGAKELEELFK------SLESLSK 541
Db   1322 IQLEEEQDAKSHADSSSRRLLAELEELKKRVAKETSDKQKAQDQRANYQRENESLKA--- 1378

Qy   --
Query Match   7.48; Score 204.5; DB 1; Length 1509;
Best Local Similarity 22.2%; Pred. No. 0.086;
Matches 127; Conservative 107; Mismatches 218; Indels 119; Gaps 25;

```
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -------------------------------------------------------------------------
DR   EMBL; M35012; AAA28713.1; -.
DR   PIR; A36014; A36014.
DR   PIR; B36014; B36014.
DR   HSSP; P08799; 1MND.
DR   FlyBase; FBgn0005634; zip.
DR   InterPro; IPR000048; IQ.
DR   InterPro; IPR002928; Myosin_tail.
DR   InterPro; IPR001609; myosin_head.
DR   Pfam; PF00612; IQ; 1.
DR   Pfam; PF00063; myosin_head; 1.
DR   Pfam; PF01576; Myosin_tail; 1.
DR   PRINTS; PR00193; MYOSINHEAVY.
DR   ProDom; PD000355; myosin_head; 1.
DR   SMART; SM00015; IQ; 1.
DR   SMART; SM00242; MYSc; 1.
DR   PROSITE; PS50096; IQ; 1.
●
KW   Myosin; Alternative splicing; Coiled coil; Actin-binding;
KW   ATP-binding; Calmodulin-binding;
FT   DOMAIN       1    829       MYOSIN HEAD-LIKE.
FT   DOMAIN     830    859       IQ.
FT   DOMAIN     886   2017       COILED COIL (POTENTIAL).
FT   NP_BIND    225    232       ATP.
FT   DOMAIN     250    260       25 KDA/50 KDA JUNCTION.
FT   DOMAIN     682    694       50 KDA/20 KDA JUNCTION.
FT   DOMAIN     705    727       ACTIN-BINDING.
FT   DOMAIN     742    758       REACTIVE SULFHYDRYL/ACTIN-BINDING.
FT   DOMAIN    1303   2017       LIGHT MEROMYOSIN (LMM).
FT   DOMAIN    1303   1970       ALPHA-HELICAL TAILPIECE.
FT   DOMAIN    1971   2017       GLOBULAR TAILPIECE (LMM).
FT   VARSPLIC     1     45       MISSING (IN SHORT ISOFORM).
SQ   SEQUENCE   2017 AA;  232016 MW;  73E3CB02BA8F2528 CRC64;
```

```
Query Match     7.0%;  Score 193.5;  DB 1;  Length 2017;
Best Local Similarity   23.3%;   Pred. No. 0.35;
Matches 131;  Conservative 87;  Mismatches 181;  Indels 163;  Gaps 27;

Qy     26 LTEISKKITDSNAVLL--------AVKEVEALLSSID------EIAAKAIGKKIHQNNGLDTEYN 76
          :||  ||                  |                  | | |      ||
Db   1136 LTQTLIRIDESATKATAQAKQRELESQLAEIQEDLEAEKARAKAEKVRD-------- 1187

Qy     77 HNGSLLAGAYAISTLIKQKDLGLKNEGLKEKIDA----------AKKCSETFINKLKEKHTD 128
          ||  |||                ||  |:|                | |:||     :|
Db   1188 ------------LSEELEALKNE-LLDSLDTTRAAQELRSKREQELATLKKSLEEET 1231

Qy    129 LGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVVLSKAAKEMLANSVKELTSPVVA 188
          :|  ||                |||
Db   1232 VNHEGVL-ADMRH-----KHSQELNSINDQLENLRK-AKTVL----------- 1266

Qy    189 ESPAMGSNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVK--EVETLVLSID 246
          :|:                                      |:|    |:
Db   1267 -EKAGTLEAEENADLATELRSVNSSRQEN--DRRRKQAESQIAELQVKLAEIERARSELQ 1323

Qy    247 ELAKKAIGQKIDNNNGLAALNNQ--NGSLLAGAYAISTLITEKLSKLKNLEELKTEIAKA 304
          : ||                     :|
Db   1324 EKCTK-LQQEAEN------IIVNQLEEAELKASA----------- 1364

Qy    305 KKCSEEFT---------NKLKSGHADLGKQDATDDHAK------AAILKTHATTDKGA 347
          ||                :|
Db   1365 QQLLEETRQKLGLSSKLRQIESEKEALQEQLEEDDEAKRNYERKLAEVTTQMQEIKKKA 1424

Qy    348 KEFKDLFESVEGLLKAAQ---VALTNSVEKLGHRNNSGGDSASTNPDESAKGPNLTVISK 404
          :| |||                |
Db   1425 EEDADLAKELEEGKKRLNKDIEALERQVKELIAQND------RLDKSKK-----KIQS 1471

Qy    405 KITDSNAFLLA-----VKEVEALLSSIDEL---SKAIGKKIKND-GTLDNEANRNESLIAG 456
          ||  |                |:|          |||
Db   1472 ELEDAATELEAQRTKVELELKEKQKNEDKILAEEKAISEQIAQERDTAEREAREKE---- 1526

Qy    457 AYEISKLITQKLSVLNSEELKKIKEAKDCSQKFTTKLKDSHAELGIGQSVQDDNAKKAIL 516
          :| ||   |      :||||       :   :     ||       :   ||    ||
Db   1527 ------------TKVLSV--SRELEDARFDKIEDLENRKKT--------------LA 1561

Qy    517 KTHGTKDKGAKELEELFKSLES 538
          :|:|  ||   ||:|:   :|||
Db   1562 NTQGTADKNVHELEKAKRALES 1583
```

```
RESULT 10
MYS2_DICDI

ID   MYS2_DICDI     STANDARD;     PRT;   2116 AA.
AC   P08799;
DT   01-NOV-1988 (Rel. 09, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   MYOSIN II HEAVY CHAIN, NON MUSCLE.
GN   MHCA.
OS   Dictyostelium discoideum (Slime mold).
OC   Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX   NCBI_TaxID=44689;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87092266; PubMed=3540939;
RA   Warrick H.M., de Lozanne A., Leinwand L.A., Spudich J.A.;
RT   "Conserved protein domains in a myosin heavy chain gene from
RT   Dictyostelium discoideum.";
RL   Proc. Natl. Acad. Sci. U.S.A. 83:9433-9437(1986).
RN   [2]
RP   PHOSPHORYLATION SITES, AND MUTAGENESIS.
RC   STRAIN=AX2;
RX   MEDLINE=90355583; PubMed=2387408;
RA   Lueck-Vielmeter D., Schleicher M., Grabatin B., Wippler J.,
RA   Gerisch G.;
RT   "Replacement of threonine residues by serine and alanine in a
RT   phosphorylatable heavy chain fragment of Dictyostelium myosin II.";
RL   FEBS Lett. 269:239-243(1990).
RN   [3]
RP   PHOSPHORYLATION SITES.
RX   MEDLINE=88112226; PubMed=2828113;
RA   Wagle G., Noegel A., Scheel J., Gerisch G.;
RT   "Phosphorylation of threonine residues on cloned fragments of the
RT   Dictyostelium myosin heavy chain.";
RL   FEBS Lett. 227:71-75(1988).
RN   [4]
RP   X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 1-762.
RX   MEDLINE=95345066; PubMed=7619795;
RA   Fisher A.J., Smith C.A., Thoden J.B., Smith R., Sutoh K., Holden H.M.,
RA   Rayment I.;
RT   "X-ray structures of the myosin motor domain of Dictyostelium
RT   discoideum complexed with MgADP.BeFx and MgADP.AlF4-.";
RL   Biochemistry 34:8960-8972(1995).
RN   [5]
RP   X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 1-762.
RX   MEDLINE=95345067; PubMed=7619796;
RA   Smith C.A., Rayment I.;
RT   "X-ray structure of the magnesium(II)-pyrophosphate complex of the
RT   truncated head of Dictyostelium discoideum myosin to 2.7-A
RT   resolution.";
RL   Biochemistry 34:8973-8981(1995).
RN   [6]
RP   X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 1-762.
RX   MEDLINE=96206189; PubMed=8611530;
RA   Smith C.A., Rayment I.;
RT   "X-ray structure of the magnesium(II).ADP.vanadate complex of the
RT   Dictyostelium discoideum myosin motor domain to 1.9-A resolution.";
RL   Biochemistry 35:5404-5417(1996).
RN   [7]
RP   X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 1-762.
RX   MEDLINE=97452580; PubMed=9305951;
RA   Gulick A.M., Bauer C.B., Thoden J.B., Rayment I.;
RT   "X-ray structures of the MgADP, MgATPgammaS, and MgAMPPNP complexes
```

```
RT   of the Dictyostelium discoideum myosin motor domain.";
RL   Biochemistry 36:11619-11628(1997).
RN   [8]
RP   X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 1-762.
RX   MEDLINE=98070605; PubMed=9405148;
RA   Bauer C.B., Kuhlman P.A., Bagshaw C.R., Rayment I.;
RT   "X-ray crystal structure and solution fluorescence characterization
RT   of Mg.2'(3')-O-(N-methylanthraniloyl) nucleotides bound to the
RT   Dictyostelium discoideum myosin motor domain.";
RL   J. Mol. Biol. 274:394-407(1997).
CC   -!- FUNCTION: MYOSIN IS A PROTEIN THAT BINDS TO ACTIN & HAS ATPASE
CC       ACTIVITY THAT IS ACTIVATED BY ACTIN.
CC   -!- SUBUNIT: MYOSIN II HEAVY CHAIN IS TWO-HEADED, IT SELF-ASSEMBLES
CC       INTO FILAMENTS. HEXAMER OF 2 HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI
CC       LIGHT CHAIN SUBUNITS (MLC) AND 2 REGULATORY LIGHT CHAIN SUBUNITS
CC       (MLC-2).
CC   -!- SUBCELLULAR LOCATION: HIGHEST CONCENTRATION IN THE POSTERIOR CELL
CC       CORTEX.
CC   -!- DOMAIN: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
CC       MEROMYOSIN (LMM) AND 1 HEAVY MEROMYOSIN (HMM). IT CAN BE FURTHER
CC       SPLIT INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
CC       SUBFRAGMENT (S2).
CC   -!- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CC       CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CC       CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC   -!- PTM: PHOSPHORYLATION INHIBITS THICK FILAMENT FORMATION AND REDUCES
CC       THE ACTIN-ACTIVATED ATPASE ACTIVITY.
CC   -!- MISCELLANEOUS: DICTYOSTELIUM MYOSIN II HAS NO K(2)EDTA ATPASE
CC       ACTIVITY, PERHAPS CORRELATED WITH THE ABSENCE OF A CYS AT THE SH-1
CC       POSITION (688).
CC   -!- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 IQ DOMAIN.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; M14628; AAA32227.1; -.
DR   PIR; A26555; A26555.
DR   PIR; S00250; S00250.
DR   PDB; 1MMA; 03-DEC-97.
DR   PDB; 1MMD; 17-AUG-96.
DR   PDB; 1MMG; 03-DEC-97.
DR   PDB; 1MMN; 03-DEC-97.
DR   PDB; 1MND; 17-AUG-96.
DR   PDB; 1MNE; 17-AUG-96.
DR   PDB; 1VOM; 23-DEC-96.
DR   PDB; 1LVK; 28-JAN-98.
DictyDb; DD01008; mhcA.
DR   InterPro; IPR000048; IQ.
DR   InterPro; IPR001609; myosin_head.
DR   Pfam; PF00612; IQ; 1.
DR   Pfam; PF00063; myosin_head; 1.
DR   PRINTS; PR00193; MYOSINHEAVY.
DR   ProDom; PD000355; myosin_head; 1.
DR   SMART; SM00015; IQ; 1.
DR   SMART; SM00242; MYSc; 1.
DR   PROSITE; PS00096; IQ; 1.
KW   Myosin; Coiled coil; ATP-binding; 3D-structure;
KW   Calmodulin-binding; Methylation; Alkylation; Phosphorylation.
FT   DOMAIN        1    761       MYOSIN HEAD-LIKE.
FT   DOMAIN      762    791       IQ.
FT   NP_BIND     179    186       ATP.
FT   DOMAIN      817   2116       COILED COIL (POTENTIAL).
FT   DOMAIN      638    660       ACTIN-BINDING.
FT   DOMAIN      738    752       ACTIN-BINDING.
FT   MOD_RES     130    130       METHYLATION (DI-) (POTENTIAL).
FT   MOD_RES     678    678       ALKYLATION (SH-1).
FT   MOD_RES    1823   1823       PHOSPHORYLATION (BY MHCK).
```

```
FT   MOD_RES    1833   1833       PHOSPHORYLATION (BY MHCK).
FT   MOD_RES    2029   2029       PHOSPHORYLATION (BY MHCK).
SQ   SEQUENCE   2116 AA;  243871 MW;  2FC3770B1EE56A1 CRC64;
```

```
Query Match             7.0%;  Score 193.5; DB 1;  Length 2116;
Best Local Similarity   20.8%;  Pred. No. 0.37;
Matches  140; Conservative  92; Mismatches 257; Indels 185; Gaps 24;

Qy      9 DGNTSANSADSSVKG-------PNLTEISKKITDSNAVLLAVKEVEALLSSSIDEIAAKAIG 62
            ||:|      ||      |      ||:| :| |:|||  ||  |  ||:  :|
Db    964 DGQSDTISRLEKIKDELQKEVEELTESFSEESKDKGVL--EKTRVRLQSELDDLTVR---- 1018

Qy     63 KKIHQNNGLDTEYNHNGSLLAGAYAISTLIRQKLDGLKNEGLKEKIDAAKKCSETFTNKL 122
              |||     | |:        |||  ||    ||| |:   |   |||    ||| ||
Db   1019 --LDSETKDKSELLROKKKLEELLKQVQEALAAE------TAAKLAQEEAANKKL 1064

Qy    123 KEKHTDLGKEGVTDADAKEAILKTNGT-------------------KTKGA---- 154
            |:||     |   |    |  ||:                         ||
Db   1065 QGEYTELNEKINSEVTARSNVEKSKKTLESQLVAVNNELDEEKKNRDALEKKKKALDAML 1124

Qy    155 EELGKLFES-----------VEVLSKAAKEMLANSVKELTSPVV------------ 187
            ||| |          :|:  |                             ---
Db   1125 EEMKDQLESTGGEKKSLIYDLVKVKQESDMEALRNQISELQSTIAKLEKIKSTLEGEVARLQ 1184

Qy    188 ----AESPAMGSNSGKGG----------DSASTNPADESAKGPNLTEISKKITDSNAFVLAVK 236
                |||   |                  | |:|:  |  |||  |  |  ||| |  |:
Db   1185 GELEAEQLAK-SNVEKQKKVELDLEDKSAQLAEETPAAKQALDKIKKKLEQE------LS 1237

Qy    237 EVETLIVLSIDELAKRAIGQKIDNNNGLAALNNQNGSLLAGAYAISTLIITEKL---SKLKN 293
            :|                ||    | |  |    |  ||  |  ||   | |      |:|
Db   1238 EVQT--QLSEANNKNVNSDSTNKHLETSFNNLKELEAEQKAKQALEKKRLGLESELKH 1294

Qy    294 LEE---------------LKTEIAKAKKCSEEFTNKLKSGHADLGKQDATDDHAKAA 335
            |:|              |:|::  |  || |  |||  |  ||||  ||:  |
Db   1295 VNEQLEEKKQKESNEKRKVDLEKEVSELKDQIEEEVASKKAVTEARNKKEESELDEIKRQ 1354

Qy    336 ILKTHAITTDKGAKEFKDL-------FESVEGLL------KAAQVALTNSVKELGHR 378
            |:|                      :||  ||        |:|    ||  |  ||
Db   1355 YADVVSSRDKSVEQLKTLQAKNEELRNTAEEABGGQLDRAERSKKRAEFDLEEAAVRNLEEE 1414

Qy    379 ---------------NNSGGDSASTNPD-ESAKGPN----------LTVISKKITDS 409
                            ||  ||   |  |             |  ||:|||:|
Db   1415 TAKKVKAEKAMKKAEFDYRSTKSELDDAKNVSSEQVYQIKRLNEELSELRSVLEEADERC 1474

Qy    410 NAFLLAVKEVEALLSSI-DELSKAIGKKIKNDGTLDNEANRNESLLIAGAYEISKLITQKL 468
            || |    |  | || |        |    ||  ||  |:|  ||  |  || | |||||
Db   1475 NSAIRAKKTAESALESLKDEIDAANNAKARAE------RKSKELEVRVAELEESLEDKS 1527

Qy    469 SVLNSEELKKIKIEARDCSQKFTTKLKDSHAELGIQSVQD-DNAKRAILKTHGTKDKGAK 527
            ::| | :    | || ||: | |  |   | :|   ||       |     |:|||   :
Db   1528 GTVNVEFIRKKDAEIDDLRARL------DRETESRIKSDEDKKNTRKQFADLEAKVEEAQR 1582

Qy    528 E---LEELFKSLES 538
            |   ||  ||||| |||
Db   1583 EVVTIDRLKKLES 1596
```

```
RESULT 11
REST_CHICK
ID   REST_CHICK    STANDARD;    PRT;   1433 AA.
AC   O42184; O42228; O57562; O57564;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   RESTIN (CYTOPLASMIC LINKER PROTEIN-170) (CLIP-170).
GN   RSN.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
```

```
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99137792; PubMed=9469933;
RA   Griparic L., Volosky J.M., Keller T.C. III;
RT   "Cloning and expression of chicken CLIP-170 and restin isoforms.";
RL   Gene 206:195-208(1998).
RN   [2]
RP   SEQUENCE OF 17-1139 FROM N.A. (CLIP-170(11) AND CLIP-170(11+35)).
RC   TISSUE=Pectoralis muscle;
RA   Griparic L., Keller T.C. III;
RT   "Identification and expression of two novel CLIP-170/Restin isoforms
RT   expressed predominantly in muscle.";
RL   Submitted (FEB-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: SEEMS TO BE A INTERMEDIATE FILAMENT ASSOCIATED PROTEIN
CC       THAT LINKS ENDOCYTIC VESICLES TO MICROTUBULES (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC, ASSOCIATED WITH THE
CC       CYTOSKELETON (BY SIMILARITY).
CC   -!- ALTERNATIVE PRODUCTS: AT LEAST 4 ISOFORMS ARE PRODUCED BY
CC       ALTERNATIVE SPLICING.
CC   -!- SIMILARITY: CONTAINS 2 CAP-GLY DOMAINS.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AF014012; AAC60344.1; -.
DR   EMBL; AF020764; AAC60345.1; -.
DR   EMBL; AF045650; AAC03547.1; -.
DR   EMBL; AF045651; AAC03548.1; -.
DR   InterPro; IPR000938; CAP-Gly.
DR   InterPro; IPR001878; ZnF_CCHC.
DR   Pfam; PF01302; CAP_Gly; 2.
DR   SMART; SM00343; ZnF_C2HC; 1.
DR   PROSITE; PS00845; CAP_GLY_1; 2.
KW   Cytoskeleton; Microtubules; Coiled coil; Alternative splicing.
FT   DOMAIN        79    121       CAP-GLY 1.
FT   DOMAIN       144    207       SER-RICH.
FT   DOMAIN       235    277       CAP-GLY 2.
FT   DOMAIN       305    332       SER-RICH.
FT   DOMAIN       351   1353       COILED COIL (POTENTIAL).
FT   DOMAIN      1414   1427       CCHC-BOX.
FT   VARSPLIC     458    458       MISSING (IN SHORT ISOFORM).
FT   VARSPLIC     458    492       TQTKLEHARIKELEEQSLLFEKTKADKLQORELEDTR ->
FT                                 RKRQISEDPEN (IN ISOFORM CLIP-170(11)).
FT   VARSPLIC     803    803       S -> GGSSKVS (IN ISOFORM CLIP-170(11)).
FT   VARSPLIC     458    458       T -> RKRQISEDPENI (IN ISOFORM CLIP-
FT                                 170(11+35)).
FT   CONFLICT     309    309       K -> R (IN REF. 2; AAC03547).
FT   CONFLICT     440    440       E -> V (IN REF. 2; AAC03548).
SQ   SEQUENCE   1433 AA;  161026 MW;  5631CE868349BE23 CRC64;
```

```
Query Match             7.0%;
Best Local Similarity  22.6%;
Matches  137; Conservative  98; Mismatches  280; Indels  92; Gaps  23;

Qy       19 ESVKGPNLTETSKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGKKIHQNNGLDTEY--- 75
            |||: |        | : | | :|| ||  |  |      |  |  :  |   ||
Db      708 ESVEDQHLVEMEDTLNKLQEAEIKVKELDVLQAKCNE-QTKLIGSLTQOIRASEEKLLDL 766

Qy       76 ------NHHG------ 75
Db      767 AALQKANSEGKLEIQKLSEQLQAAEKOIQNLETEKVSNLTKELQGKEQKLLDLEKNLSA- 825

Qy      118 FTNKLK---EKHTDGKEGVTDA---------DAKEAILKTNGTKTRKGA----EELGKLFE 162
            ||:|        | |  | |               ||  |  | | |        | ||
Db      826 -VNQVKDSLEKELQLLKEKFTSAVDGAENAQORAMQETINKLLNQKEEQFALMSSELEQLKS 884

Qy      163 SVEVLSKAAKEMLANSVKELTSPVVAESPAMGSNSGKGGDSASTNPADES--AKGPNLTE 220
            ::|:  ||  |      |      |:          :   |        :   |: ||
```

```
RESULT 12
MYSB_CAEEL

ID   MYSB_CAEEL      STANDARD;    PRT;  1966 AA.
AC   P02566;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   MYOSIN HEAVY CHAIN B (MHC B).
GN   UNC-54 OR MYO-4.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=83273600; PubMed=6576334;
RA   Karn J., Brenner S., Barnett L.;
RT   "Protein structural domains in the Caenorhabditis elegans unc-54
RT   myosin heavy chain gene are not separated by introns.";
RL   Proc. Natl. Acad. Sci. U.S.A. 80:4253-4257(1983).
RN   [2]
RP   SEQUENCE OF 850-1966 FROM N.A.
RX   MEDLINE=82272395; PubMed=7202124;
RA   McLachlan A.D., Karn J.;
RT   "Periodic charge distributions in the myosin rod amino acid sequence
RT   match cross-bridge spacings in muscle.";
RL   Nature 299:226-231(1982).
RN   [3]
RP   SEQUENCE OF 1876-1966 FROM N.A.
RX   MEDLINE=83238892; PubMed=6571695;
RA   Wills N., Gesteland R.F., Karn J., Barnett L., Bolten S.,
RA   Waterston R.H.;
RT   "The genes sup-7 X and sup-5 III of C. elegans suppress amber
RT   nonsense mutations via altered transfer RNA.";
RL   Cell 33:575-583(1983).
CC   -!- FUNCTION: MUSCLE CONTRACTION.
CC   -!- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
CC       HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
CC       AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
CC   -!- SUBCELLULAR LOCATION: THICK FILAMENTS OF THE MYOFIBRILS.
CC   -!- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CC       CYCLES OF A 28-RESIDUE TAIL SEQUENCE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
```

```
Db      885 NLTVMETKLKEREREQLTEAKVKLENDIAEIMKSSGDSSAQLMKMNDELRLKERQLEQ 944

Qy      221 ISKKITDSNAFVLAVK-------------EVETLVLSIDELAKRAIGKIDNNGLAALN 267
            :|: | | |  : ::              |||  | | :|    | : ||     |
Db      945 IQLELTKANEKAVQLQKNVEQTAQKAEQSQQEFLKTHQPEEL-KKMQDQLTDMKKQMETSQ 1003

Qy      268 NQNGSLLAG-AYAISTLITEKLSKLKNLEELKTEIAKAKKCSEEFTNKLKSGHADLGKQD 326
            |  |  |   | |: | |  | |  |  | | |  | :| |      | |   ||  | | |
Db     1004 NQYKDLQARKYEKEISEMITKHDADIKGFKQNLLDAEEALKAAQOKNDELETQAEELKQ- 1062

Qy      327 ATDDHAKRAAILKTHATTDKGAKEFKDLFESVEGLLKAAQOVALTNSVKELGHRNNSGGDSA 386
            | |  | |  |  | | |   |  |  |  | |   | | ||  :  | :| :|
Db     1063 --AEQAKA---------DKRAEE----VLQTMEKVTKEKDAIHQEKIETLASLENSRQINE 1108

Qy      387 STNPDESAKGPNLIVISKKITDSNAFL-LAVKEVEALLSSIDELSKAIGKKIKNDGTLDN 445
            |:  |    ||  :|  : | | | |    | |   | | :|    | :| | |   | |
Db     1109 KLQNELDMLKQNNLKNEEELTKSKELLNLENKVVEELKKEFEALKLAAAQKSQOLAALQE 1168

Qy      446 E-ANRNESLLIAGAYEIS---KLITQKLSVLNSE-------ELKKKIKEAKDCSQKFTTK 493
             : | || ::|   ::      | :| | |::         |||:|   |     :
Db     1169 ENVKLAEELGRSRDEVTSHQKLEEERSVLNNQLLEMKKRESTLKKEIDDERASLQK---S 1225

Qy      494 LKDSHAEELGIQSVQDDNAKKAILKTHGTRDKGAKELEELFKSLESLSKAAQOAALTNSVKE 553
            | |:|  |  || | || | | |  |||  |  | | | | :| : | |   |   | |
Db     1226 ISDTSALITQKDEELEKLRNEITVLRG-ENASAKTLQSVVKTLES----DKLKLEEKVKN 1280

Qy      554 LTNPVVA 560
            | ||
Db     1281 LEQKLKA 1287
```

```
CC  -!- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
CC      ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
CC  -!- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
CC      MEROMYOSIN (LMM) AND 1 HEAVY MEROMYOSIN (HMM). IT CAN LATER BE
CC      SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
CC      SUBFRAGMENT (S2).
CC  -!- MISCELLANEOUS: THERE ARE FOUR DIFFERENT MYOSIN HEAVY CHAINS IN C.
CC      ELEGANS.
CC  -!- MISCELLANEOUS: MHC A AND MHC B ARE FOUND EXCLUSIVELY IN THE BODY
CC      WALL MUSCLE. THEY CO-ASSEMBLE INTO BODY WALL THICK FILAMENT.
CC  -!- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC  -----------------------------------------------------------------------
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----------------------------------------------------------------------
DR  EMBL; J01050; AAA28124.1; -.
DR  EMBL; V01494; CAA24738.1; -.
DR  PIR; A02992; MWKW.
DR  HSSP; P08799; 1MND.
DR  InterPro; IPR002928; Myosin_tail.
DR  InterPro; IPR001609; myosin_head.
DR  Pfam; PF00063; myosin_head; 1.
DR  Pfam; PF01576; Myosin_tail; 1.
DR  PRINTS; PR00193; MYOSINHEAVY.
DR  ProDom; PD000193; Myosin_head; 1.
DR  SMART; SM00242; MYSc; 1.
KW  Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
KW  ATP-binding; Methylation; Alkylation; Multigene family.
FT  DOMAIN        1     850       MYOSIN HEAD-LIKE.
FT  DOMAIN      851    1966       COILED COIL (POTENTIAL).
FT  DOMAIN      851    1164       ALPHA-HELICAL TAILPIECE (S2).
FT  DOMAIN     1165    1176       HINGE.
FT  DOMAIN     1165    1966       LIGHT MEROMYOSIN (LMM).
FT  NP_BIND     177     184       ATP (BY SIMILARITY).
FT  DOMAIN      665     687       ACTIN-BINDING.
FT  DOMAIN      769     783       ACTIN-BINDING.
FT  MOD_RES     128     128       METHYLATION (TRI-) (POTENTIAL).
FT  MOD_RES     705     705       ALKYLATION (SH-1).
FT  MOD_RES     715     715       ALKYLATION (SH-2).
FT  CONFLICT   1337    1337       E -> L (IN REF. 2).
FT  CONFLICT   1880    1880       I -> L (IN REF. 2).
SEQUENCE   1966 AA;  225125 MW;  B66F0BB2FE27B67F CRC64;
```

```
Query Match             7.0%;
Best Local Similarity   21.5%;
Matches   127;  Conservative  104;

Qy    16 SADESVKGPNLTEISKKITDSNAVLLAVK----EVEALLSSIDEI--AAKAIGKKIHQ-NN 69
            | |::||:| || ::| |  |||: | :: |    ||:| |:   ||| :  : :::||
Db  1358 SLEEEIEGKN--EILRQLSKANADIQQWKARFEGEGLLKA-DELEDAKRRQAQKINELQE 1414

Qy    70 GLDTEYNHNGSL-LAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNKLKEKHTD 128
            ||  |  ||  | | ||| |:| | |||  :| | || | :| | | : : |  | |
Db  1415 ALDAANSKNASLEKTKSRLVGDLDDAQVDVERANGVASALEKKQKGFDKIIDEWRKKTDD 1474

Qy   129 LGKEGYTDADAKEAILKTNGT------KTKGAEELGKLFESVEVLSKAAKEMLANSVKELTSP 185
            : |  ||:| : |  | |        : | :|   |:  | | || ::|| :|| ||:|||
Db  1475 LAAE------LDGAQRDLRNTSDLFRAKNAQE--ELAEVVEGLRRENKS-LSQEIKDLTDQ 1527

Qy   186 VVAESPAMGSNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETLVLSI 245
            :|  |:|  |: | :|  |    | ||  | ||:|:| |  :: :: : : |:
Db  1528 L----------------------------------------GEGRS-- 1545

Qy   246 DELAKKAIGQKIDNNGLAALNNQNGSLLAGAYAISTLITEKLSKLKNLEELKTEIAK-A 304
            |||::|||| ||:|:||  | : || | ||| |:|||::|| | |:||  | |::|| |
Db  1546 -EIEKEELQHALDEAE---AALEAEESKVLRAQVEVS---------QIRSEIEKRI 1588
```

```
Qy   305 KKCSEEFTNKLKSGHADLGKDGATTDDHAKA------AILKTHATTDKGAKEFKDLFESVEG 359
            ||| |: |  |:| |||| :| :|   |        | |:| |  ||| ::| : :  ||
Db  1589 QEKEEFENTRK------------NHARALESMQASLETEA---KGKAELLRIKKKLEG 1632

Qy   360 LIKAAQVALTNSVKELGHRNNSGGGDSAST--NPDESAKGPNLTVISKKITDSNA---FLL 414
            |:|| ||  |  :|  :  :: :   || |    | :: ||||   |:| |:|:||    :|
Db  1633 DINELEIA-------LDHANKANADAQKNLKRYQEQVRELQLQVEEQRNGADTREQFFN 1685

Qy   415 AVKEVEALLSSIDEL--SKAIGKKIKNDGTLD------NEANRNESLIAGA----- 457
            |:|:: || :| |    || :|||:| |  |        |  :| ::  |
Db  1686 AEKRATLLQSEKEELLVANEAAERARKQAEYEAADARDQANEANAQVSSLTSAKRKLEGE 1745

Qy   505 SVQDDNAKRAILKTHGTKDKGAKELEELFKSLESLSKAQQAALTNSVKEL 554
            :|:|:|||  :|  | | :|  | |||| ::  || | |::||:|  :|
Db  1746 IQAIHADLDETLNEYKAAEEERSKKAIADATRLAEELRQEQEHSQHVDRLRKGLEQQIKEI 1805

Db  1806 QVRLDEAEAAALK--GGKKVIAK-LEQRVRELEESELDGEQRRFQDANKNL 1852
```

```
Score 192;  DB 1;  Length 1966;
Pred. No. 0.39;  Indels 146;  Gaps 28;
Mismatches 213;
```

```
Query Match      6.9%;  Score 189;  DB 1;  Length 1022;
Best Local Similarity 23.5%;  Pred. No. 0.25;  Mismatches 204;  Indels 220;  Gaps 37;
Matches 158;  Conservative 89;

Qy   5 NSGKDGNTSANSADE--SVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAA----  58
         :  :  |    : ||||    | |  ::|| ||:
Db 322 NHGQSVALSQTIAEDLTHVQGPS-HETHKPIIIPN------QELESSIEQHTSQQVP 371

Qy  59 -----KAIGKKIHQNNGLDTEYNHNGSL---LAGAYAISTLIKQKLDGL----KNEG 103
            |||  ||| : |   ::| || : ||  |
Db 372 PITTFNKSLQPKISQIHQLQPQQAQSSGIPNPVLNAANALSTSMQDLLNNINSYLTKNQD 431

Qy 104 LKEKIDAAKKCSETFNKLK------EKHTDL-------GKRGVTADA---------139
        :::| || || :|   ::   |   |
Db 432 INKQSDLIKEAAIAILNNKKSDFAEKQYNIIDLAKNIFSNKDIIADAKVNVNTLLETIQ 491

Qy 140 ------KEAILKTNGTKTIKGAEELGKLFESVEVLSKAAKEMLANSVKELTSPVVAES 190
             | |:|   |  | :| |
Db 492 NDQNTLDIKKSKILEDIVAIYLNSENI-ELKQKQQILEKVVD--IGLSIKDISRVVAVD 548

Qy 191 PAMGNSNGKGGDSASTNPADESAKGPNLT----------EIS---KKITDSNAFVLAVKEV 238
        ||||||||||||||||||||||||||||              :||  |  :||| ||
Db 549 SIMDT------VIKSNIANEDKEKIFITVFDQINSYEFSNVAKQKLLDS----ILKKTA 597

Qy 239 ETLVLSIDELAKKAIGQKIDN----------NNGL------AALNNQNGSLLAGAY 278
         | | |:|| ||| ||:|:||          ||||      ::||||||||||||
Db 598 ETQVLSPEQ--QQLMNQLDNITTEHYKRDTIEKVNNILLEPLSNFALKTTNIQVMTSNV 655

Qy 279 AISTLITEKLSKL--------KNLEEL-------KTEIAKAKKCSEEFTNKLKSGHADLGKQD 326
          |||  || ::|        |::||:       :||:||||||||||:||||||
Db 656 LDSPVQIEMKSRLIQVVTKTVAESALVEPKDKTEIVKG--------IGKTI 698

Qy 327 ATDD------HAKAAILKTHATTDKGAKEFKDLFESVEGLKAAQV--ALTNSVKELGHR 378
         ||:       | ||:|||   | |  | :|| || ||  ||
Db 699 VTHSDTLSPLHDKVVIMGSVA--------KGIVESKNDLLDRELIIAGLVDGIYE---- 745

Qy 379 NNSGGDSASTNPDESAKGPNLTV--ISKKITDSNAFLLAVKEVEALLSSIDELS-KAIGK 435
        ||||||||||||||||||||||   ||||||||||  ||||||||||||||:|  ||||
Db 746 ------AKGDNAVVHAISSMIANSN---INQSEKEALKRSQDVVSEKVLDK 787

Qy 436 KIKNDGTLDNE---ANRNESLIAG-------AYEISKLITQKLSVLN------SEE 475
         |:| :  |:|   |  :|   |         || ||:||| | :|        |||
Db 788 EION---LDRELKAQNINESKLHDDIYNKTQDVANALKNVII---TVLDDNSGQRGVSEE 841

Qy 476 LKKKIKE-AKDCSQKFTTKLKDSHAELGIQSVQDDNAKKAILKT-HGTKDKGAKELEELF 533
         |||:|:  | |||:|   || |:||| || :  |||| |:|| |   | :||| |||
Db 842 APKKVSSLLNDISKRIEKINNLRAMLS----QDGN----LKTFEEKKDEATKKVDELV 892

Qy 534 KSLESLSKAAQ 544
         | : | ||
Db 893 KAFDNKSSTEE 903
```

```
RESULT 14
REST_HUMAN
ID   REST_HUMAN              STANDARD;      PRT;   1427 AA.
AC   P30622;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-APR-1993 (Rel. 25, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   RESTIN (CYTOPLASMIC LINKER PROTEIN-170 ALPHA-2) (CLIP-170) (REED-
DE   STERNBERG INTERMEDIATE FILAMENT ASSOCIATED PROTEIN).
GN   RSN.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Peripheral blood monocytes;
RX   MEDLINE=92289675; PubMed=1600942;
RA   Bilbe G., Delabie J., Brueggen J., Richener H., Asselbergs F.A.M.,
RA   Cerletti N., Sorg C., Odink K., Tarcsay L., Wiesendanger W.,
RA   de Wolf-Peeters C., Shipman R.;
RT   "Restin: a novel intermediate filament-associated protein highly
RT   expressed in the Reed-Sternberg cells of Hodgkin's disease.";
RL   EMBO J. 11:2103-2113(1992).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92405160; PubMed=1356075;
RA   Pierre P., Scheel J., Richard J.E., Kreis T.E.;
RT   "CLIP-170 links endocytic vesicles to microtubules.";
RL   Cell 70:887-900(1992).
CC   -!- FUNCTION: SEEMS TO BE A INTERMEDIATE FILAMENT ASSOCIATED PROTEIN
CC       THAT LINKS ENDOCYTIC VESICLES TO MICROTUBULES.
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC, ASSOCIATED WITH THE
CC       CYTOSKELETON.
CC   -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
CC       SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN THE REED-STERNBERG CELLS
CC       OF HODGKIN'S DISEASE.
CC   -!- SIMILARITY: CONTAINS 2 CAP-GLY DOMAINS.
```

```
DR   EMBL; X64838; CAA46050.1; -.
DR   EMBL; M97501; AAA35693.1; -.
DR   PIR; S22695; S22695.
MIM; 179838;
DR   InterPro; IPR000938; CAP-Gly.
DR   InterPro; IPR001878; Znf_CCHC.
Pfam; PF01302; CAP_GLY; 2.
SMART; SM00343; ZnF_C2HC; 1.
PROSITE; PS00845; CAP_GLY; 2.
PROSITE; PS00845; CAP_GLY; 1.
KW   Cytoskeleton; Microtubules; Coiled coil; Alternative splicing.
FT   DOMAIN     60    125   CAP-GLY 1.
FT   DOMAIN    143    204   SER-RICH.
FT   DOMAIN    214    279   SER-RICH.
FT   DOMAIN    304    331   SER-RICH.
FT   DOMAIN    350   1342   COILED COIL (POTENTIAL).
FT   DOMAIN   1408   1421   CCHC-BOX.
FT   VARSPLIC  457    491   MISSING (IN SHORT ISOFORM).
FT   CONFLICT 1069   1069   D -> E (IN REF. 2).
SQ   SEQUENCE  1427 AA;  160989 MW;  0A4F166DD94254E8 CRC64;
```

```
Query Match      6.9%;  Score 189;  DB 1;  Length 1427;
Best Local Similarity 21.6%;  Pred. No. 0.36;  Mismatches 249;  Indels 140;  Gaps 26;
Matches 135;  Conservative 100;

Qy  28 EISKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGKKIHQNNGLD------TEVNHNGS 80
        :|:| :| | :  ||  | :|||:|||   |  |  | |||| |          | |:|
Db 523 ESNKPAGDVDMSLSLIQEISSLQEKLE-----VIRTDHQREITSLKEHFGAREETHQKE 576

Qy  81 LLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFNKLKEK--HTDLGKEGVTADA 139
         :| |  || ||:  : |  ::| |:|:||||||||||||||:|  | ||||
Db 577 IKALYTATEKLSKE-----NESLKSKLEHANKENSDVIALWKSLETATASHQQAMEEL 630

Qy 140 KEAILKTNGTKTIKGAEELGKLFESVEVLSKAAKEMLANSVKELTSPVVAESPAMGSN-- 196
         | :| | | ::  |  | : | : : : |  |||:| :  :|         ||  | |
Db 631 KVSFSKGLGTETAEFAELKTQIEKMRLDYQHEIENLQNQQDSERAAHAKEMALRAKLMK 690

Qy 197 --SGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETLVLSIDELAK--- 250
          ::|   : | ||  |  | || :| ||::|  | | :| |::| || |:| ||
Db 691 VIKEKENSLEAIRSKLDKAEDQHLVEMEDLNKLQEAEIKVKELEVLQAKCNEQTKVIDN 750

Qy 251 -----KAIGQKI------DNNNGLAALNNQNGSLLAGAYAISTL--------I 284
            ||:| :|          |::||||||||||||||||||        |
Db 691 VIKEKENSLEAIRSKLDKAEDQHLVEMEDLNKLQEAEIKVKELEVLQAKCNEQTKVIDN 750
```

```
RESULT 15
ZIP1_YEAST
ID   ZIP1_YEAST     STANDARD;      PRT;    875 AA.
AC   P31111;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   SYNAPTONEMAL COMPLEX PROTEIN ZIP1.
GN   ZIP1 OR YDR285W OR D9819.9.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BR1824-3B;
RX   MEDLINE=93161412; PubMed=7916652;
RA   Sym M., Engebrecht J.A., Roeder G.S.;
RT   "ZIP1 is a synaptonemal complex protein required for meiotic
RT   chromosome synapsis.";
RL   Cell 72:365-378(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288C / AB972;
RA   Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z.,
RA   Favello A., Fulton L., Gattung S., Greco T., Kirsten J.,
RA   Kucaba T., Hallsworth K., Hawkins J., Hillier L., Jier M.,
RA   Johnson D., Johnston L., Langston Y., Latreille P., Le T.,
RA   Mardis E., Menezes S., Miller N., Nhan M., Pauley A., Peluso D.,
RA   Rifken L., Riles L., Taich A., Trevaskis E., Vignati D.,
RA   Wilcox L., Wohldman P., Vaudin M., Wilson R., Waterston R.;
RL   Submitted (MAR-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: REQUIRED FOR MEIOTIC CHROMOSOME SYNAPSIS AND CELL CYCLE
CC       PROGRESSION. MAY ACT AS A MOLECULAR ZIPPER TO BRING HOMOLOGOUS
CC       CHROMOSOMES IN CLOSE APPOSITION. ZIP1 MAY ENCODE THE TRANSVERSE
CC       FILAMENTS OF THE SYNAPTONEMAL COMPLEX.
CC   -!- SUBCELLULAR LOCATION: SYNAPSED MEIOTIC CHROMOSOMES.
```

```
DR   EMBL; L06487; AAA35239.1; -.
DR   EMBL; U51031; AAB64474.1; -.
DR   PIR; A45173; A45173.
DR   PIR; S30868; S30868.
DR   SGD; S0002693; ZIP1.
KW   Nuclear protein; Meiosis; Coiled coil.
FT   DOMAIN      177    333    COILED COIL (POTENTIAL).
FT   DOMAIN      397    438    COILED COIL (POTENTIAL).
FT   DOMAIN      456    752    COILED COIL (POTENTIAL).
FT   CONFLICT     55     55    T -> A (IN REF. 1).
SQ   SEQUENCE  875 AA;  100035 MW;  674F12625CD9DFD CRC64;
```

```
Query Match    6.8%;  Score 187;  DB 1;  Length 875;
Best Local Similarity 20.1%;  Pred. No. 0.25;
Matches 115;  Conservative 117;  Mismatches 197;  Indels 142;  Gaps 25;
```

```
Db   751 FTSQLKATEEKLLDDALRKASSEGKSEMKKLRQOLEAAAEKQIKHLEIEKNAESSKASSI  810
Qy   285 TEKLS----KLKNLEELKTEIAKAKKCSEEFTNKLKSGHADLGKQDATDDHA-KAAILKT  339
Db   811 TRELQGRELKLINLQENLSEVSQVKEPLEKELQILKEKFAEASEEAVSVQRSMQETVNKL  870
Qy   340 HA--------TTDKGAKEFKDLFESVEGLLKAAQVAALTNSVKELGHRNNS  381
Db   871 HQKEEQFNMLSSDLKLRENLADMEAK-FREKDEREEQLIKAKE-KLENDIAEI---MKM  925
Qy   382 GGDSAS------TNPDESAKGPNLTVISKKITDSNAFLLAVKEVEALLSSIDELSKAIGKKI  437
Db   926 SGDNSSQLTKMNDELRLKERDVEELQLKLTKAN-----ENASFLQKSIEDMT-----V  973
Qy   438 KNDGTLDNEANRNESLIAGAYEISKLITQKLSVLNSEELKKKI------KEAKDCSQKF  490
Db   974 KAEQSQQEAAKHE------EEKKELERKLS------DLEKKMETSHNQCQELKARYERA 1021
Qy   491 TTKLKDSHAELGIQSVQDDNAKKAILKTHGTKDKDGAK------ELEELFKSLESLSKA  542
Db  1022 TSETKTKHEEI-LQNLQ-----KTLLDTE-DKLKGAREENSGLLQELEELRKQADK-AKA 1073
Qy   543 AQAA--------LTNSVKELTNPVVA  560
Db  1074 AQTAEDAMQIMEQMTKEKTEFLAS 1097
```

```
Qy    20 SVKGPNLT--EISKKITDSNAVLLAV-KEVEALLSSIDEIAAKAIGKKIHQNGLDTEYN   76
Db   274 SIENLNKTINDLGKNKEADAELMKKGKEIEYLKRELDCGSQOLSEEKI-KNSSLIQEMG  332
Qy    77 HNGSLLAGAYAISTLIKQKLDLGLKNEGLLEKI-------DAAKCSEEFTNKLKEKHTDL  129
Db   333 KNR----------EEMIKSIENFFSEDKAHHLLQFNKFEERVHDLFEKKL-QKHFDV  378
Qy   130 GKEGYTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKELTSPVVAE  189
Db   379 AKDTLNVGLRNTTVELSSNTETMLKQQYEDIKENLEQKMSSSKDEMAKTINELS--VTQK  436
Qy   190 SPAMGSNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETLVLSIDELA  249
Db   437 GLIMG------------------------------VQEELLTSSG-----NIQTALVS--  459
Qy   250 KKAIGQKIDNNNGLAAALNNQNGSLLAGAYAISTLIITEKLSKLKNL-EELKTEIAKAKCS  308
Db   460 ----------EMNNTRQELLDDA------SQTAKNYASLENLVKAKVKAEIVQSNE-Y  499
Qy   309 EEFTNKLKSGHADLGKQDATDDHAKAAILKTHAPTDKGAKE--FKDLFESVEGLLKAAQV  366
Db   500 EERIKHLESERSTLSSQ--------KNQIISSLGTKEAQYEDLVKKLEA--KNIEI  545
Qy   367 A-LTNSVKELGHRNNSGGDSASTNPDESAKGPNLTVISK----KITDSNAFLLAVKEVE  420
Db   546 SQISGKEQSLTEKNENLSNELKVQODQLEKLNNLNITTKSNYENKISSQNEI-----VK  599
Qy   421 ALLSSIDELSKAIGKKI------KNDGTLDNEA--NRNESLIAGAYEISKLITQKLSVLN  472
Db   600 ALVSENDTLKQRIQQLVEIKENEQKDHTTKLEAFQKNNEQLQKLNVEVVVQLKAHEL---  655
Qy   473 SEELKKKIKEAKDCSQKFTKLKDSHAELGI--QSV-------QDDNAKKAILKTHGTKD  523
Db   656 --ELEEQNRHLKNCLEKKETGVEESLSDVKTLKQQVIVLKSEKQDIPAEKLELQDN---  709
Qy   524 KGAKELEELFKSLESLSKRAAQVAALTNSVKEL  554
Db   710 --LESLEEVTKNLQQKVQSQKRELEQKIKEL  738
```

Search completed: March 18, 2002, 10:12:03
Job time: 981 sec

THIS PAGE BLANK (USPTO)

OM protein - protein search, using sw model

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

Run on: March 18, 2002, 10:10:55 ; Search time 124.19 Seconds
(without alignments)
659.574 Million cell updates/sec

Title: US-09-596-746a-52
Perfect score: 2750
Sequence: 1 MACNNSGKDGNTSANSADES..........KAAQAALTNSVKELTNPVVA 560

Scoring table: BLOSUM62
Searched: 473505 seqs, 146272329 residues
Gapop 10.0 , Gapext 0.5

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_17:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	926	33.7	192	2	Q9S3P3	Q9s3p3 borrelia bu
2	921	33.5	192	2	Q9R7B1	Q9r7b1 borrelia bu
3	910	33.1	191	2	P70818	P70818 borrelia bu
4	908	33.0	193	2	Q9RR53	Q9rr53 borrelia bu
5	893	32.5	200	2	Q9RQR9	Q9rqr9 borrelia bu
6	886	32.2	194	2	Q9S3P4	Q9s3p4 borrelia bu
7	885	32.2	182	2	Q9R7B2	Q9r7b2 borrelia bu
8	884	32.1	203	2	P96507	P96507 borrelia af
9	881	32.0	207	2	O07336	O07336 borrelia bu
10	880	32.0	212	2	Q9KIM4	Q9kim4 borrelia bu
11	878	31.9	207	2	O45187	O45187 borrelia ga
12	878	31.9	207	2	Q9REF8	Q9ref8 borrelia ga
13	878	31.9	207	2	Q9REF7	Q9ref7 borrelia ga
14	877	31.9	212	2	Q44669	Q44669 borrelia af
15	876	31.9	203	2	O50623	O50623 borrelia af
16	874	31.8	200	2	O08139	O08139 borrelia bu
17	853	31.0	177	2	Q44999	Q44999 borrelia bu
18	846	30.8	179	2	Q44987	Q44987 borrelia af
19	844	30.7	179	2	Q44988	Q44988 borrelia af
20	843.5	30.7	178	2	Q9REH7	Q9reh7 borrelia bu
21	834.5	30.3	180	2	Q44991	Q44991 borrelia af
22	832	30.3	179	2	Q44986	Q44986 borrelia af
23	829	30.1	175	2	Q9R7B4	Q9r7b4 borrelia bu
24	805	29.3	189	2	P94224	P94224 borrelia bu
25	802	29.2	163	2	Q9R7A9	Q9r7a9 borrelia bu
26	798	29.0	207	2	O45177	O45177 borrelia ga
27	794	28.9	200	2	Q44996	Q44996 borrelia ga
28	786	28.6	207	2	Q49581	Q49581 borrelia ga
29	773	28.1	159	2	Q9R7B0	Q9r7b0 borrelia bu
30	759.5	27.6	211	2	Q44720	Q44720 borrelia af
31	757.5	27.5	193	2	P94237	P94237 borrelia bu
32	756.5	27.5	199	2	Q9S3P2	Q9s3p2 borrelia bu
33	755.5	27.5	206	2	P96509	P96509 borrelia af
34	745.5	27.1	204	2	P96512	P96512 borrelia af
35	705.5	25.7	206	2	Q44719	Q44719 borrelia bu
36	700	25.5	192	2	Q9S3P2	Q9s3p2 borrelia bu
37	698	25.4	193	2	Q9S3P2	Q9s3p2 borrelia bu
38	696.5	25.3	192	2	P94234	P94234 borrelia bu
39	691	25.1	193	2	P94234	P94234 borrelia bu
40	687.5	25.0	204	2	Q44989	Q44989 borrelia af
41	683.5	24.9	211	2	Q49576	Q49576 borrelia bu
42	678.5	24.7	194	2	O31122	O31122 borrelia af
43	678.5	24.7	211	2	O31123	O31123 borrelia bu
44	677	24.6	212	2	Q9KIM5	Q9kim5 borrelia af
45	676.5	24.6	209	2	Q44671	Q44671 borrelia af

ALIGNMENTS

RESULT 1

Q9S3P3

ID   Q9S3P3   PRELIMINARY;   PRT;   192 AA.
AC   Q9S3P3;
DT   01-MAY-2000 (TREMBLrel. 13, Created)
DT   01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C (FRAGMENT).
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=IP2;
RX   MEDLINE=96296448; PubMed=8709845;
RA   Livey I., Gibbs C.P., Schuster R., Dorner F.;
RT   "Evidence for lateral transfer and recombination in OspC variation in
RT   Lyme disease Borrelia.";
RL   Mol. Microbiol. 18:257-269(1995).
DR   EMBL; L42887; AAB36995.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
FT   NON_TER   1
FT   NON_TER   192
SQ   SEQUENCE   192 AA; 20287 MW; 11846F7AC84C7E3D CRC64;

Query Match   33.7%; Score 926; DB 2; Length 192;
Best Local Similarity   99.5%;
Matches 188; Conservative   0; Mismatches   1; Indels   0; Gaps   0;

Qy   3 CNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAIG 62
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 CNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAIG 60

Qy   63 KKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNKL 122
              ||||||||||||| |||||||||||||||||||||||| | ||||||| |||| ||||||
Db   61 KKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKREKIDAARKCSEFTNKL 120

Qy   123 KEKHTDLGHKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKEL 182

```
Db  121 KEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKEL 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  183 TSPVVAESP 191
        |||||||||
Db  181 TSPVVAESP 189
```

RESULT 2
Q9R7B1

```
ID   Q9R7B1       PRELIMINARY;   PRT;   192 AA.
AC   Q9R7B1;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C (FRAGMENT).
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TETS;
RX   MEDLINE=94478003; PubMed=9336916;
RA   Ras N.M., Postic D., Foretz M., Baranton G.;
RT   "Borrelia burgdorferi sensu stricto, a bacterial species 'made in the
RT   U.S.A.'?";
RL   Int. J. Syst. Bacteriol. 47:1112-1117(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TETS;
RA   Marti-Ras N., Postic D., Foretz M., Baranton G.;
RL   Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U91798; AAB81895.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
FT   NON_TER       1
FT   NON_TER     192
SQ   SEQUENCE   192 AA;  20297 MW;  67705024A20AAF764 CRC64;

Query Match             33.5%;  Score 921;  DB 2;  Length 192;
Best Local Similarity   98.4%;
Matches  187;  Conservative    2;  Mismatches   1;  Indels   0;  Gaps   0;

Qy    1 MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3 ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA 62

Qy   61 IGKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN 120
         |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db   63 IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN 122

Qy  121 KLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  123 KLKEKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVK 182

Qy  181 ELTSPVVAES 190
         ||||||||||
Db  183 ELTSPVVAES 192
```

RESULT 3
P70818

```
ID   P70818       PRELIMINARY;   PRT;   191 AA.
AC   P70818;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN (FRAGMENT).
OSPC.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=2-1498 CA4;
RA   Probert W.S., Crawford M.R., Cadiz R.B., LeFebvre R.B.;
RL   Submitted (SEP-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; L81131; AAB06569.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
FT   NON_TER       1
SQ   SEQUENCE   191 AA;  20126 MW;  D2B9B1C82B4DC3C0 CRC64;

Query Match             33.1%;  Score 910;  DB 2;  Length 191;
Best Local Similarity   98.9%;  Pred. No. 2.1e-34;
Matches  186;  Conservative    0;  Mismatches   2;  Indels   0;  Gaps   0;

Qy    4 NNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGK 63
         ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 NNSGKGGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKAIGK 60

Qy   64 KIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNKLK 123
         ||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 KIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTNKLK 120

Qy  124 EKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKELT 183
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 EKHTDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKELT 180

Qy  184 SPVVAESP 191
         ||||||||
Db  181 SPVVAESP 188
```

RESULT 4
Q9RR53

```
ID   Q9RR53       PRELIMINARY;   PRT;   193 AA.
AC   Q9RR53;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C (FRAGMENT).
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=OC1;
RX   MEDLINE=99091544; PubMed=9872945;
RA   Wang I.N., Dykhuizen D.E., Qiu W., Dunn J.J., Bosler E.M., Luft B.J.;
RT   "Genetic diversity of ospC in a local population of Borrelia
RT   burgdorferi sensu stricto.";
RL   Genetics 151:15-30(1999).
DR   EMBL; AF029860; AAB86543.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
FT   NON_TER       1
FT   NON_TER     193
SQ   SEQUENCE   193 AA;  20502 MW;  5EFDB5AFF8986D1E CRC64;

Query Match             33.0%;  Score 908;  DB 2;  Length 193;
Best Local Similarity   98.4%;  Pred. No. 2.6e-34;
Matches  184;  Conservative    2;  Mismatches   1;  Indels   0;  Gaps   0;

Qy    1 MACNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA 60
         :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7 ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA 66

Qy   61 IGKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSETFTN 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    67  IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAARKCSETFTN  126
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   121  KLKEKHTDLGKEGVTDADAKEAILKTNGTKTKTRGAEELGKLFESVEVLSKAAKEMLANSVK  180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   127  KLKEKHTDLGKEGVTDADAKEAILKTNGTKTKTRGAEELGKLFESVEVLSKAAKEMLANSVK  186

Qy   181  ELTSPVV  187
             |||||||
Db   187  ELTSPVV  193


RESULT 5

ID   Q9RQR9              PRELIMINARY;        PRT;    200 AA.
AC   Q9RQR9;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C PRECURSOR (FRAGMENT).
OS   Borrelia burgdorferi (Lyme disease spirochete).
OG   Plasmid cp26.
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=2E7;
RX   MEDLINE=20002545; PubMed=10531219;
RA   Hofmeister E.K., Glass G.E., Childs J.E., Persing D.H.;
RT   "Population dynamics of a naturally occurring heterogeneous mixture of
RT   Borrelia burgdorferi clones.";
RL   Infect. Immun. 67:5709-5716(1999).
DR   EMBL; AF074464; AAD23911.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
FT   NON_TER        1      1
FT   NON_TER      200    200
SQ   SEQUENCE   200 AA;  21307 MW;  A22F158758BEBB6B CRC64;

Query Match             32.5%;  Score 893;  DB 2;  Length 200;
Best Local Similarity   98.4%;  Pred. No. 1.3e-33;
Matches  181;  Conservative   2;  Mismatches   1;  Indels   0;  Gaps   0;

Qy    1  MACNNSGKDGNTSANSADESVKGPNLTEISKRITDSNAVLLAVKEVEALLSSIDEIAAKA   60
            ::|||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db   17  ISCNNSGKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEALLSSIDEIAAKA   76

Qy   61  IGKKIHQNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAARKCSETFTN  120
            ||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db   77  IGKKIHQNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAARKCSETFTN  136

Qy  121  KLKEKHTDLGKEGVTDADAKEAILKTNGTKTKTRGAEELGKLFESVEVLSKAAKEMLANSVK  180
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  137  KLKEKHTDLGKEGVTDADAKEAILKTNGTKTKTRGAEELGKLFESVEVLSKAAKEMLANSVK  196

Qy  181  ELTS  184
            ||||
Db  197  ELTS  200


RESULT 6

ID   Q9S3P4              PRELIMINARY;        PRT;    194 AA.
AC   Q9S3P4;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C (FRAGMENT).
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JSB;
RX   MEDLINE=96296448; PubMed=8709845;
RA   Livey I., Gibbs C.P., Schuster R., Dorner F.;
RT   "Evidence for lateral transfer and recombination in OspC variation in
RT   Lyme disease Borrelia.";
RL   Mol. Microbiol. 18:257-269(1995).
DR   EMBL; L42883; AAB36991.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
FT   NON_TER        1      1
FT   NON_TER      194    194
SQ   SEQUENCE   194 AA;  20446 MW;  CEEDC9FA5DF0D68F CRC64;

Query Match             32.2%;  Score 886;  DB 2;  Length 194;
Best Local Similarity   90.7%;  Pred. No. 2.5e-33;
Matches  186;  Conservative   2;  Mismatches   3;  Indels  14;  Gaps   2;

Qy  195  SNSGKGGDSASTNPADESAKGPNLTEISKRITDSNAFVLAVKEVETLVLSIDELAKKAIG  254
            :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2  NNSGKGGDSASTNPADESAKGPNLTEISKRITDSNAFVLAVKEVETLVLSIDELAKKAIG   61

Qy  255  QKIDNNNGLAALNNQNGSLLAGAYAISTLITEKLSKLKNLEELKTEIAKAKKCSEEFTNK  314
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   62  QKIDNNNGLAALNNQNGSLLAGAYAISTLITEKLSKLKNLEELKTEIAKAKKCSEEFTNK  121

Qy  315  LKSGHADLGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGLLKAAQVALTNSVKE  374
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  122  LKSGHADLGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGLLKAAQVALTNSVKE  181

Qy  375  LGHRNNSGGDSASTNP--DESAKGP  397
            |:|              :|    |  |
Db  182  L---------------TSPVVAESPKKP  194


RESULT 7

ID   Q9R7B2              PRELIMINARY;        PRT;    182 AA.
AC   Q9R7B2;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C (FRAGMENT).
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TETS;
RX   MEDLINE=97478003; PubMed=9356916;
RA   Marti-Ras N., Postic D., Foretz M., Baranton G.;
RT   "Borrelia burgdorferi sensu stricto, a bacterial species 'made in the
RT   U.S.A.'?";
RL   Int. J. Syst. Bacteriol. 47:1112-1117(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TETS;
RA   Marti-Ras N., Postic D., Foretz M., Baranton G.;
RL   Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U91797; AAB11894.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
FT   NON_TER        1      1
FT   NON_TER      182    182
SQ   SEQUENCE   182 AA;  19202 MW;  422146F99AA57BF2 CRC64;
```

```
Qy    7  GKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEVALLSSIDEIAAKAIGKKIH   66
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  GKDGNTSANSADESVKGPNLTEISKKITDSNAVLLAVKEVEVALLSSIDEIAAKAIGKKIH   60

Qy   67  QNNGLDTEYNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSEFTFNKLKEKH   126
         ||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  QNNGLDTENNHNGSLLAGAYAISTLIKQKLDGLKNEGLKEKIDAAKKCSEFTFNKLKEKH   120

Qy  127  TDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKELTSPV  186
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121  TDLGKEGVTDADAKEAILKTNGTKTKGAEELGKLFESVEVLSKAAKEMLANSVKELTSPV  180

Qy  187  VA  188
         ||
Db  181  VA  182
```

```
RESULT 8
ID   P96507      PRELIMINARY;      PRT;   203 AA.
AC   P96507;
DT   01-MAY-1997 (TREMBLrel. 03, Created)
DT   01-MAY-1997 (TREMBLrel. 03, Last sequence update)
DT   01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C (FRAGMENT).
OS   Borrelia afzelii.
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=29518;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BFOX;
RA   Masuzawa T., Komikado T., Fukui T., Yanagihara Y.;
RL   Submitted (JAN-1997) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PBO;
RA   Masuzawa T., Sawaki K., Yoshii T.;
RL   Submitted (JAN-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB003345; BAA19078.1; -.
DR   EMBL; AB009895; BAA24125.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
FT   NON_TER        1      1
FT   NON_TER      203    203
SQ   SEQUENCE   203 AA;  21549 MW;  94A885FBBACFE1C4 CRC64;

Query Match             32.1%;  Score 884;  DB 2;  Length 203;
Best Local Similarity   99.4%;
Matches  180;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy  195  SNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVEVLVLSIDELAKKAIG  254
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   14  NNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVEVLVLSIDELAKKAIG   73

Qy  255  QKIDNNNGLLAALNNQNGSLLAGAYAISTLITEKLSKLKNLEELKTEIAKAKKCSEEFTNK  314
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   74  QKIDNNNGLLAALNNQNGSLLAGAYAISTLITEKLSKLKNLEELKTEIAKAKKCSEEFTNK  133

Qy  315  LKSGHADLGKQODATDDHAKAAILKTHATTDKGAKEFKDLFESVEGLLKAAQVALTNSVKE  374
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  134  LKSGHADLGKQODATDDHAKAAILKTHATTDKGAKEFKDLFESVEGLLKAAQVALTNSVKE  193

Qy  375  L  375
         |
Db  194  L  194
```

```
RESULT 9
ID   Q07336      PRELIMINARY;      PRT;   207 AA.
AC   Q07336;
DT   01-NOV-1998 (TREMBLrel. 08, Created)
DT   01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT   01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PBi;
RX   MEDLINE=93268136; PubMed=8098841;
RA   Jauris-Heipke S., Fuchs R., Motz M., Preac-Mursic V., Schwab E.,
RA   Will G., Wilske B.;
RT   "Genetic heterogeneity of the genes coding for the outer surface
RT   protein C (OspC) and the flagellin of Borrelia burgdorferi.";
RL   Med. Microbiol. Immunol. 182:37-50(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PBi;
RX   MEDLINE=93239332; PubMed=8478108;
RA   Wilske B., Preac-Mursic V., Jauris S., Pradel I., Soutschek E.,
RA   Schwab E., Wanner G.;
RT   "Immunological and molecular polymorphisms of OspC, an immunodominant
RT   major outer surface protein of Borrelia burgdorferi.";
RL   Infect. Immun. 61:2182-2191(1993).
DR   EMBL; X69595; CAA49305.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
SQ   SEQUENCE   207 AA;  22320 MW;  4A525ABDAE63B5C8 CRC64;

Query Match             32.0%;  Score 881;  DB 2;  Length 207;
Best Local Similarity   99.5%;  Pred. No. 4.6e-33;
Matches  181;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy  379  NNSGGDSASTNPDESAKGPNLTVISKKITDSNAFLLAVKEVEVALLSSIDELSKRAIGKKIK  438
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   20  NNSGGDSASTNPDESAKGPNLTVISKKITDSNAFLLAVKEVEVALLSSIDELSKRAIGKKIK   79

Qy  439  NDGTLDNEANRNESLIAGAYEISKLIITQKLSVLNSEELKKKIKEAKDCSQKFTTKLKDSH  498
         ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db   80  NDGTLDNEANRNESLIAGAYEISKLIITQKLSVLNSEELKEKIKEAKDCSQKFTTKLKDSH  139

Qy  499  AELGIOSVQDDNAKKAILKTHGTKDKGAKELELFKSLESLSKRAAQAALTNSVKELTNPV  558
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  140  AELGIOSVQDDNAKKAILKTHGTKDKGAKELELFKSLESLSKAAQAALTNSVKELTNPV  199

Qy  559  VA  560
         ||
Db  200  VA  201
```

```
RESULT 10
ID   Q9KIM4      PRELIMINARY;      PRT;   212 AA.
AC   Q9KIM4;
DT   01-OCT-2000 (TREMBLrel. 15, Created)
DT   01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT   01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C.
OS   Borrelia afzelii.
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=29518;
RN   [1]
RP   SEQUENCE FROM N.A.
```

RC   STRAIN=PGAU;
RA   Dykhuizen D.E., Guttman D.S., Luft B.J.;
RT   "Antigenic variation and intragenic recombination in the ospc gene of
RT   Borrelia burgdorferi.";
RL   Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF230185; AAF75626.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
SQ   SEQUENCE 212 AA; 22495 MW; C51DAF803BF2E7D4 CRC64;

Query Match   32.0%; Score 880; DB 2; Length 212;
Best Local Similarity   90.2%; Pred. No. 5.2e-33;
Matches 185; Conservative   2; Mismatches   4; Indels 14; Gaps   2;

Qy   195 SNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETLVLSIDELAKKAIG 254
             ||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   20 NNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETLVLSIDELAKKAIG 79

Qy   255 QKIDNNNGLAALNNQNGSLLAGAYAISTLITEKLSKLKNLEELKTETAKAKKCSEEFTNK 314
             |||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db   80 QKIDNNNGLAALNNQNGSLLAGAYAISTLITEKLSKLKNLEELKTEIAKAKKCSEEFTNK 139

Qy   315 LKSGHADLGKQDATDDHAKAAILKTHATTDKGAKEKDLFESVEGLLKAAQVALTNSVKE   374
             |||||||||||||||||||||||| ||||||||||| ||||||||||||||||||||||
Db   140 LKSGHADLGKQDATDDHAKAAILKPHATTDKGAKEFKDLFESVEGLLKAAQVALTNSVKE   199

Qy   375 LGHRNNSGGDSASTNP--DESAKGP 397
             :|: ||||| || |
Db   200 L----------------TSPVVAESPKP 212

RESULT 11
ID   Q45187
AC   Q45187;
DT   01-NOV-1996 (TREMBLrel. 01, Created)
DT   01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT   01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C PRECURSOR.
OS   Borrelia garinii.
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=29519;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PTROB;
RX   MEDLINE=95213332; PubMed=7699024;
RA   Wilske B., Jauris-Heipke S., Lobentanzer R., Pradel I.,
RA   Preac-Mursic V., Rossler D., Soutschek E., Johnson R.C.;
RT   "Phenotypic analysis of outer surface protein C (OspC) of Borrelia
RT   burgdorferi sensu lato by monoclonal antibodies: relationship to
RT   genospecies and OspA serotype.";
RL   J. Clin. Microbiol. 33:103-109(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PTROB;
RX   MEDLINE=95395018; PubMed=7665660;
RA   Jauris-Heipke S., Liegl G., Preac-Mursic V., Rossler D.,
RA   Soutschek E., Will G., Wilske B.;
RT   "Molecular analysis of genes encoding outer surface protein C (OspC) of
RT   Borrelia burgdorferi sensu lato: relationship to ospA genotype and
RT   evidence of lateral gene exchange of ospC.";
RL   J. Clin. Microbiol. 33:1860-1866(1995).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PSCF;
RA   Marconi R.T., Hohenberger S., Jauris-Heipke S., Schulte-Spechtel U.,
RA   Lavole C.P., Roessler D., Wilske B.;
RT   "Genetic analysis of B.garinii OspA-serotype 4 strains associated with
RT   neuroborreliosis: evidence for extensive genetic homogeneity.";
RL   Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.

DR   EMBL; X83554; CAA58544.1; -.
DR   EMBL; AJ236908; CAB46238.1; -.
DR   EMBL; AJ132793; CAB46231.1; -.
DR   EMBL; AJ132796; CAB46234.1; -.
DR   EMBL; AJ132797; CAB46235.1; -.
DR   EMBL; AJ132798; CAB46236.1; -.
DR   EMBL; AJ236907; CAB46237.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
FT   SIGNAL   1   18   POTENTIAL.
FT   CHAIN   19   207   POTENTIAL.
SQ   SEQUENCE 207 AA; 22321 MW; 20089AA76E63A49D CRC64;

Query Match   31.9%; Score 878; DB 2; Length 207;
Best Local Similarity   98.9%; Pred. No. 6.2e-33;
Matches 180; Conservative   2; Mismatches   0; Indels   0; Gaps   0;

Qy   379 NNSGGDSASTNPDESAKGPNLTVISKKITDSNAFLLAVKEVEALLSSIDELSKAIGKKIK 438
             |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db   20 NNSGGDSASTNPDESAKGPNLTVISKKITDSNAFLLAVKEVEALLSSIDELSKAIGKKIK 79

Qy   439 NDGTLDNEANRNESLIAGAYEISKLITTQKLSVLNSEELKKKIKEAKDCSQKFTTKLKDSH 498
             ||||||||||||||||||||||||||||||||||||||| ||||||||| ||||||||||
Db   80 NDGTLDNEANRNESLIAGAYEISKLITTQKLSVLNSEELKEKIKEAKDCSEKFTTKLKDSH 139

Qy   499 AELGIQSVQDDNAKKAILKTHGTKDKGAKELEELFKSLESLSKAAQAALTNSVKELTNPV 558
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   140 AELGIQSVQDDNAKKAILKTHGTKDKGAKELEELFKSLESLSKAAQAALTNSVKELTNPV 199

Qy   559 VA 560
             ||
Db   200 VA 201

RESULT 12
ID   Q9REF8
AC   Q9REF8;
DT   01-MAY-2000 (TREMBLrel. 13, Created)
DT   01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C (FRAGMENT).
OS   Borrelia garinii.
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=29519;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PWA;
RX   MEDLINE=20033636; PubMed=10565915;
RA   Marconi R.T., Hohenberger S., Jauris-Heipke S., Schulte-Spechtel U.,
RA   Lavoie C.P., Roessler D., Wilske B.;
RT   "Genetic analysis of Borrelia garinii OspA serotype 4 strains
RT   associated with neuroborreliosis: evidence for extensive genetic
RT   homogeneity.";
RL   J. Clin. Microbiol. 37:3965-3970(1999).
DR   EMBL; AJ132794; CAB46232.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
FT   NON_TER   207   207
SQ   SEQUENCE 207 AA; 22366 MW; 30CFEAA6D913A49D CRC64;

Query Match   31.9%; Score 878; DB 2; Length 207;
Best Local Similarity   98.9%; Pred. No. 6.2e-33;
Matches 180; Conservative   2; Mismatches   0; Indels   0; Gaps   0;

Qy   379 NNSGGDSASTNPDESAKGPNLTVISKKITDSNAFLLAVKEVEALLSSIDELSKAIGKKIK 438
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 13

```
Q9REF7
ID   Q9REF7        PRELIMINARY;      PRT;    207 AA.
AC   Q9REF7;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN.
GN   OSPC.
OS   Borrelia garinii.
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=29519;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20033636; PubMed=10565915;
RA   Marconi R.T., Hohenberger S., Jauris-Heipke S., Schulte-Spechtel U.,
RA   LaVoie C.P., Roessler D., Wilske B.;
RT   "Genetic analysis of Borrelia garinii OspA serotype 4 strains
RT   associated with neuroborreliosis evidence for extensive genetic
RT   homogeneity.";
RL   J. Clin. Microbiol. 37:3965-3970(1999).
DR   EMBL; AJ132795; CAB46233.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
SQ   SEQUENCE   207 AA;   22249 MW;   20889AA76E60C49D CRC64;
```

```
Query Match             31.9%;    Score 878;   DB 2;    Length 207;
Best Local Similarity   98.9%;
Matches   180;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   379 NNSGGDSASTNPDESAKGPNLTVISKKITDSNAFLLAVKEVEALLSSIDELSKAIGKKIK 438
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    20 NNSGGDSASTNPDESAKGPNLTVISKKITDSNAFLLAVKEVEALLSSIDELSKAIGKKIK 79

Qy   439 NDGTLDNEANRNESLIAGAYEISKLITQKLSVLNSEELKKKIKEAKDCSQKFTTKLKDSH 498
             |||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db    80 NDGTLDNEANRNESLIAGAYEISKLITQKLSVLNSEELKKKIKEAKDCSEKFTTKLKDSH 139

Qy   499 AELGIQSVQDDNAKKAILKTHGTKDKGAKELEELFKSLESLSKAAQAALTNSVKELTNPV 558
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   140 AELGIQSVQDDNAKKAILKTHGTKDKGAKELEELFKSLESLSKAAQAALTNSVKELTNPV 199

Qy   559 VA 560
             ||
Db   200 VA 201
```

RESULT 14

```
Q44669
ID   Q44669        PRELIMINARY;      PRT;    212 AA.
AC   Q44669;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C.
GN   OSPC.
OS   Borrelia afzelii.
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=29518;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PBO;
RX   MEDLINE=95395018; PubMed=7665660;
RA   Jauris-Heipke S., Liegl G., Preac-Mursic V., Roessler D., Schwab E.,
RA   Soutschek E., Will G., Wilske B.;
RT   "Molecular analysis of genes encoding outer surface protein C (OspC)
RT   of Borrelia burgdorferi sensu lato: relationship to ospA genotype and
RT   evidence of lateral gene exchange of ospC.";
RL   J. Clin. Microbiol 33:1860-1866(1995).
DR   EMBL; X81521; CAA57241.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
SQ   SEQUENCE   212 AA;   22514 MW;   C6EDC33BFD488DDE CRC64;
```

```
Query Match             31.9%;    Score 877;   DB 2;    Length 212;
Best Local Similarity   89.8%;
Matches   184;  Conservative    4;  Mismatches    3;  Indels   14;  Gaps    2;

Qy   195 SNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETLVLSIDELAKKAIG 254
          :|||||||||||||:|:|||||||||||||||||||||||||||||||||||||||||
Db    20 NNSGGDSASTNPDESAKGPNLTVISKKITDSNAFLLAVKEVEALLSSIDELSKAIGKKIK 79

Qy   255 QKIDNNNGLAALNNQNGSLLAGAYAISTLITEKLSKLKNLEELKTEIAKAKKCSEEFTNK 314
          :||||||||||||||||||:||:|||||||||||||||||||||||||||||||||||||
Db    80 QKIDNNNGLAALNNQNGSLLAGAYAISTLITEKLSKLKNLEELKTEIAKAKKCSEEFTNK 139

Qy   315 LKSGHADLGKQDATDDHAKAAILKTHATTDKGAKEFKDLFESVEGLLKAAQVALTNSVKE 374
          |||||||||||||| || ||||||||||| || |||||||||||||||||||||||||||
Db   140 LKSGHADLGKQDATDEHAKAAILKTHATTDKGAKEFKDLFESVEGLLKAAQVALTNSVKE 199

Qy   375 LGHRNNSGGDSASTNP--DESAKGP 397
          |:|              ||||||||
Db   200 L-------------------TSPVVAESPKKP 212
```

RESULT 15

```
O50623
ID   O50623        PRELIMINARY;      PRT;    203 AA.
AC   O50623;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C (FRAGMENT).
GN   OSPC.
OS   Borrelia afzelii.
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=29518;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PWUD;
RA   Masuzawa T., Sawaki K., Yoshii T.;
RL   Submitted (DEC-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB009899; BAA24129.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
FT   NON_TER       1      1
FT   NON_TER     203    203
SQ   SEQUENCE   203 AA;   21505 MW;   66DFF78EBACAE219 CRC64;
```

```
Query Match             31.9%;    Score 876;   DB 2;    Length 203;
Best Local Similarity   98.9%;
Matches   179;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy   195 SNSGKGGDSASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETLVLSIDELAKKAIG 254
```

```
Db    14  NNSGKGGASASTNPADESAKGPNLTEISKKITDSNAFVLAVKEVETLVLSIDELAKKAIG    73
          :||||||  |||||||||||||||||||||||||||||||||||||||||||||||||||

Qy   255  QKIDNNNGLAALNNONGSLLAGAYAISTLITTEKLSKLKNLEELKTEIAKAKKCSEEFTNK   314
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    74  QKIDNNNGLAALNNONGSLLAGAYAISTLITTEKLSKLKNLEELKTEIAKAKKCSEEFTNK   133

Qy   315  LKSGHADLGKQDADDDHAKANAILKTHATTDKGAKEFKDLFESVEGLLKAAQVALTNSVKE   374
          |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db   134  LKSGHADLGKQDATDDDHAKAAILKTHATTDKGAKEFKDLFESVEGLLKAAQVALTNSVKE   193

Qy   375  L   375
          |
Db   194  L   194
```

Search completed: March 18, 2002, 10:10:56
Job time: 979 sec

THIS PAGE BLANK (USPTO)